

UNCLASSIFIED

AD NUMBER

AD869396

LIMITATION CHANGES

TO:

Approved for public release; distribution is unlimited. Document partially illegible.

FROM:

Distribution authorized to U.S. Gov't. agencies and their contractors; Critical Technology; JAN 1970. Other requests shall be referred to Rome Air Development Center, Attn: RADC/EMNRR, Griffiss AFB, NY 13440. Document partially illegible. This document contains export-controlled technical data.

AUTHORITY

RADC ltr dtd 20 Jun 1973

THIS PAGE IS UNCLASSIFIED

AD 869396

RADC-TR-89-356, Volume II  
Technical Report  
January 1970



**MAINTAINABILITY, PREDICTION AND DEMONSTRATION TECHNIQUES**

Volume II - Maintainability Demonstration

ARINC Research Corporation

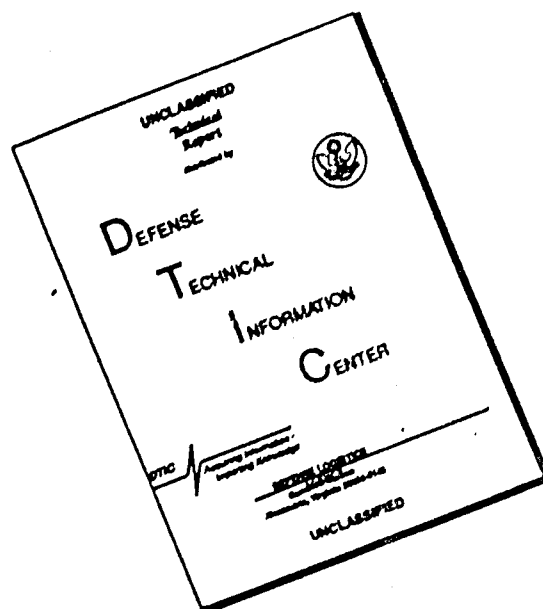
This document is subject to special  
rights reserved and may be used  
in limited quantities for internal  
use only. No other reproduction  
without permission of RADC is  
permitted.



When the Government is required to provide for the needs of the people, it is not its duty to provide for the needs of the people in a way that is not in the best interests of the people. It is the duty of the Government to provide for the needs of the people in a way that is in the best interests of the people. It is the duty of the Government to provide for the needs of the people in a way that is in the best interests of the people.

RECEIVED BY	
DATE	TIME
10	10
RECEIVED BY	
DATE	TIME
10	10

# DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

# **MAINTAINABILITY PREDICTION AND DEMONSTRATION TECHNIQUES**

## **Volume II - Maintainability Demonstration**

**Harold Balaban**

**ARINC Research Corporation**

This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of RADC (EMNRR), GAFB, NY 13440

# FOREWORD

This interim report, in two volumes, was prepared by ARINC Research Corporation, 2551 Riva Road, Annapolis, MD 21401, under Contract Number F30602-68-C-0047, Project 5519, Task 551901. The authors of Volume I were: G. Griswold; B. Retterer; A. Mitsopoulos; S. Bishop and J. Martin; of Volume II, Harold S. Balaban. ARINC's report number is 573-01-1-992.

The distribution of this document is limited by the U.S. Mutual Security Acts of 1949.

This technical report has been reviewed and is approved.

Approved:

*Eugene Fiorentino*

EUGENE FIORENTINO  
Reliability Engineering Section  
Reliability Branch

Approved:

*Joseph D. Naresky*

JOSEPH D. NARESKY  
Chief, Reliability & Compatibility Division

## ABSTRACT

Maintainability demonstration is a testing procedure for assuring the acquisition of equipment and systems with satisfactory maintainability characteristics. The results of a study to improve maintainability-demonstration procedures for Air Force equipment are presented in this report.

An industry- and Government-wide survey was conducted to provide insight into the current status of maintainability demonstration and to initiate research into the managerial, administrative, and technical aspects of demonstration. Specific recommendations and guidelines were developed on the following:

- Management planning for maintainability demonstration
- Maintainability-index selection
- Maintenance-task sampling procedures
- Statistical maintainability-demonstration test plans
- Test administration and implementation

The use of prior information for specifying numerical requirements, designing statistical sampling procedures, developing test criteria, and applying Bayesian tests was also investigated; applicable procedures and data are included in this report.

## CONTENTS

	<u>Page</u>
SECTION I: INTRODUCTION	1
1.1 The Concept of Maintainability Demonstration	1
1.2 Summary of Major Objectives	2
SECTION II: SURVEY OF MAINTAINABILITY DEMONSTRATION	5
2.1 Introduction	5
2.2 Respondents	5
2.3 Summary of Results	5
2.4 Discussion of Results	5
2.4.1 Respondent Background	6
2.4.2 Rejection Experience	6
2.4.3 Purpose of Maintainability Demonstration	6
2.4.4 Views on Aspects of Maintainability Standards or Contracts	18
2.4.5 Use of Prior Information	18
2.4.6 Difficulties Experienced	19
2.4.7 Opinions on Specified Maintainability Values	19
2.4.8 Approach to Maintainability Specifications	19
2.4.9 Lognormal-Parameter Specification	21
2.4.10 Type of Statistical Test	21
2.4.11 Views on MIL-STD-471	21
2.4.12 Approval of Stratified Sampling	22
SECTION III: DEMONSTRATION AS A MAINTAINABILITY-PROGRAM ELEMENT	23
3.1 General	23
3.2 Management Responsibilities in Maintainability Testing	24
3.2.1 Development-Test and Demonstration-Test Management	24
3.2.2 Overall Test Planning	24
3.3 The Maintainability-Demonstration-Test Plan	24
3.3.1 Responsibility for Preparing Test Plan	27
3.3.2 Contractor's Responsibility	28
3.3.3 Maintainability-Demonstration-Plan Milestones	28
3.4 Reliability and Maintainability Improvement Through Demonstration	31

## CONTENTS (continued)

	<u>Page</u>
SECTION IV: THE STATISTICAL BASIS OF MAINTAINABILITY- DEMONSTRATION	35
4.1 General	35
4.2 Decision Errors	35
4.3 The Test Hypothesis	36
4.3.1 Specification in Terms of Confidence Level	37
4.4 Relationship of Risks to Sample Size	38
4.5 The Operating-Characteristics (O.C.) Curve	38
4.6 Types of Maintainability Decision Tests	40
4.6.1 Type of Measurement	40
4.6.2 Type of Maintenance-Task-Sample Selection	40
4.6.3 Single, Multiple, and Sequential Sampling	40
4.6.4 Parametric and Nonparametric Tests	43
4.6.5 Classical and Bayesian Tests	46
SECTION V: MAINTAINABILITY-DEMONSTRATION-TEST SPECIFICATION	49
5.1 General	49
5.2 Type of Specified Maintainability Index	50
5.2.1 Type of Maintenance Action	51
5.2.2 Statistical Measures	52
5.2.3 Type of Time Measurement	54
5.2.4 Guidelines for Index Selection	59
5.3 Specified Values for the Maintainability Index	63
5.3.1 Three Basic Criteria	63
5.3.2 Specification Based on Higher-System-Level Requirements	64
5.3.3 Realism of Specified Values	74
5.3.4 Applicability of Requirements to the Demonstration Environment	84
5.4 Risk Assignment	87
5.4.1 General	87
5.4.2 Use of Prior Information in Risk Trade-off	88
SECTION VI: SELECTION OF MAINTENANCE-TASK SAMPLE	95
6.1 General	95
6.2 Natural Versus Induced Failures	95

## CONTENTS (continued)

	<u>Page</u>
6.3 Fault-Inducement Procedures	96
6.4 Design of Maintenance-Task Sample	98
6.4.1 Sample Design for a Natural-Failures Test	99
6.4.2 Sample Design for a Fault-Inducement Test	100
 SECTION VII: STATISTICAL MAINTAINABILITY-DEMONSTRATION PLANS	 121
7.1 General	121
7.2 MIL-STD-471 Plans	121
7.2.1 General	121
7.2.2 Test Method 1	122
7.2.3 Test Method 2	123
7.2.4 Test Method 3	124
7.2.5 Test Method 4	125
7.2.6 Test Method 5	125
7.2.7 Test Method 6	126
7.3 Alternative Non-Bayesian Tests	127
7.3.1 General	127
7.3.2 Comments on Sampling Procedure and Sample Size	127
7.3.3 Lognormal Distribution	127
7.3.4 Notation	127
7.3.5 Prior Estimation of $\sigma^2$	128
7.3.6 Fixed-Sample-Size Tests, Lognormal Distribution	137
7.3.7 Sequential Tests-Lognormal Distribution	154
7.3.8 Nonparametric Tests	160
7.4 Guidelines for Test Selection - Non-Bayesian Tests	184
7.4.1 Introduction	184
7.4.2 Summary of the Fourteen Tests	184
7.4.3 Decision Tree for Selecting a Test	184
7.4.4 Combinations of Two Tests	186
7.5 Bayesian Tests	187
7.5.1 General	187
7.5.2 Basic Assumptions	189
7.5.3 Maintainability Index and Test Requirements	190
7.5.4 General Bayesian Formulation of the Test	192
7.5.5 Derivation of the Posterior Density	193
7.5.6 Derivation of Cumulative Posterior Probabilities	195

## CONTENTS (continued)

	<u>Page</u>
7.5.7 Maintainability-Demonstration-Test Application	198
7.5.8 Example	201
7.5.9 Prior-Distribution Analysis	209
References for Section VII	215
SECTION VIII: TEST ADMINISTRATION AND IMPLEMENTATION	217
8.1 General	217
8.2 Test Scheduling	217
8.3 Prior-Information Requirements	219
8.4 Maintenance-Personnel Selection	219
8.5 Demonstration-Review Team	221
8.6 Demonstration-Data Forms and Records	222
8.7 Predemonstration Activities	226
8.7.1 Site Visit	226
8.7.2 Training and Indoctrination	226
8.7.3 Preparation of Task-Analysis Package	226
8.8 Conduct of the Test	227
8.9 Data Reduction and Analysis	230
8.10 Test Reporting	230
8.11 Actions Following Rejection	231
SECTION IX: CONCLUSIONS AND RECOMMENDATIONS	233
9.1 General	233
9.2 The Concept of Maintainability Demonstration	233
9.3 The Current Status of Maintainability Demonstration	233
9.4 Demonstration as a Program Element	233
9.5 The Statistical Basis of Demonstration	234
9.6 Maintainability-Demonstration-Test Specification	234
9.7 Selection of Maintenance-Task Sample	236
9.8 Statistical Maintainability-Demonstration-Test Plans	237
9.9 Test Administration and Implementation	238
APPENDIX A: THE LOGNORMAL DISTRIBUTION	A-1
APPENDIX B: MULTIPLE-REGRESSION COMPUTER PRINTOUTS FOR VARIANCE PREDICTION	B-1

## SECTION I

### INTRODUCTION

#### 1.1 THE CONCEPT OF MAINTAINABILITY DEMONSTRATION

Maintainability demonstration is the process by which a customer determines if a product he intends to buy will exhibit satisfactory maintainability characteristics. The specific approach used can range from reliance on the producer's assurance and good reputation to an extensive controlled field test of the product.

Neither of these extremes is satisfactory. The producer's reputation is more pertinent to selecting him initially, and his assurance cannot be accepted unless there are adequate factual data to support it. Often, for novel or complex equipment, it is not possible to assure that the equipment will perform satisfactorily in the field maintenance environment unless some form of controlled testing is performed. On the other hand, extensive field tests are generally costly and time-consuming.

The current Air Force policy is generally to perform limited controlled tests, employing standard statistical procedures to determine conformance with specified maintainability characteristics.

Two military standards provide the requirement and direction for conducting a maintainability-demonstration test. MIL-STD-470, "Maintainability Program Requirements" (Systems and Equipments), 21 March 1966, contains the following detailed requirement for such a test:

"Para. 5.11 Demonstrate Achievement of Maintainability Requirements

The achievement of maintainability requirements shall be demonstrated as specified in the contract. The demonstration will normally be accomplished in accordance with MIL-STD-471, 'Maintainability Demonstration', which includes contractor preparation and submission of a demonstration plan and report to the procuring activity. The demonstration plan must be responsive to the maintainability program established by the requirements of this standard. Maintainability demonstration efforts shall be integrated with other system testing requirements such as proof of design, bread-board, prototype, environmental, production and acceptance. Maintainability demonstration data will be used to incrementally verify the achievement of maintainability design requirements and to update the

maintainability parameter values from the maintainability analyses and predictions. The formal maintainability demonstration performed to determine contract compliance shall be conducted in an operational or simulated operational environment as specified in the contract."

The referenced standard, MIL-STD-471, "Maintainability Demonstration", 15 February 1966, Notice 1, 4 April 1968, provides the detailed procedures for planning and conducting maintainability-demonstration tests, and is the document most often invoked in current contracts that require such tests.

The demonstration procedure is essentially a statistical test of a hypothesis. In this case, the hypothesis is generally of the form that a specified maintainability characteristic (e.g., mean active-corrective-maintenance time) meets a specified numerical value. Accordingly, the standard approach has been one of acceptance sampling, in which known risks of wrong decisions (rejecting a satisfactory product or accepting an unsatisfactory product) are considered in relation to sample size (e.g., number of maintenance actions observed). One goal is to arrive at risk levels and sample requirements that meet existing or implied constraints.

The statistical nature of a maintainability-demonstration test imposes requirements on several aspects of the test procedure, including the test environment, the sampling procedure, and the analysis of test results. Therefore, a demonstration test cannot be evaluated or a new one proposed without careful consideration of these factors as they relate to the inferential nature of the statistical test.

Therefore, the demonstration procedure can be viewed as comprising two major areas -- the planning, management, and implementation of the test; and the statistical procedures to be employed in the decision-making process.

## 1.2 SUMMARY OF MAJOR OBJECTIVES

The overall objective of the maintainability-demonstration phase of the program to develop maintainability techniques is to develop improved procedures for planning, implementing, and evaluating maintainability-demonstration tests. Accordingly, effort was directed at both the managerial/administrative aspects and the technical and statistical aspects of the demonstration element in the maintainability-program plan.

Specific study was devoted to the following subjects:

- Management of the maintainability-demonstration program

- Prior information approaches in test design and analysis
- Statistical analysis of maintainability-demonstration test results

In the area of management and administration of the maintainability-demonstration effort, guidelines and specific approaches were developed to provide assurance that (1) the specified maintainability index is appropriate and realistic, (2) the task-sampling procedure is adequate, (3) the test is conducted in an unbiased and meaningful manner, and (4) the results are analyzed and interpreted properly.

The use of prior information, such as previous history on similar items, results of previous tests, and inputs from the maintainability-prediction efforts, were incorporated whenever possible into procedures for specification, task sampling, and test design. In particular, new Bayesian approaches for maintainability demonstration were developed.

In this report the statistical basis for demonstration is reviewed, and guidelines for selecting general approaches (e.g., fixed versus sequential sampling) are presented. An extensive set of statistical procedures for demonstration, offering improvements in rigor, efficiency, or applicability over current procedures, is described. Guidelines are presented for selecting the appropriate plan to fit particular circumstances.

## SECTION II

### SURVEY ON MAINTAINABILITY DEMONSTRATION

#### 2.1 INTRODUCTION

Although maintainability demonstration has not been applied to the same extent as reliability demonstration, it was believed that enough experience has been accumulated to warrant a survey of industry and Government personnel concerned with it. Accordingly, a comprehensive questionnaire was developed to cover both management and technical aspects of maintainability demonstration. Exhibit 1 is a copy of this questionnaire. A rather comprehensive discussion of the results of this survey is presented here because these results provide a good summary of the current status of maintainability demonstration. Particular attention is given to problems that the respondents believe have not yet been solved by current procedures.

#### 2.2 RESPONDENTS

This questionnaire was mailed to approximately 200 people and agencies actively engaged in the field of maintainability management and engineering. Four major sources were used to obtain respondents:

- The EIA-G-42 maintainability committee mailing list
- The military and Government organizations and personnel involved in MIL-STD-471 coordination
- Organizations that deal with the RADC Maintainability Section
- A list of authors on maintainability, obtained through a literature search of various technical journals and symposium proceedings

#### 2.3 SUMMARY OF RESULTS

A total of 39 usable replies was received (approximately a 25-percent return), some of which contained detailed comments on one or more aspects. Table I presents a summary tabulation of the responses received.

#### 2.4 DISCUSSION OF RESULTS

The results of the survey are discussed in this subsection. A summary of the most pertinent statistics of Table I is presented and interpreted. Indications of how improvements can and

should be made are also presented, but detailed consideration of such means is reserved for later subsections. In a sense, then, the discussions of the survey questions serve as a summary of some of the major areas to be considered in subsequent sections of this report.

#### 2.4.1 Respondent Background

Of the 39 responses received, approximately 75 percent represented industry viewpoints; approximately 75 percent of the respondents have been involved with one or more maintainability demonstrations, the total of such demonstrations exceeding 40.

The most commonly used tests were those of MIL-M23313(9) and MIL-M26512(8), which correspond to Test Methods 3 (4 applications) and 2 (9 applications), respectively, of MIL-STD-471. Thus, of 41 identified tests used for demonstration, 17 demonstrations were based on the procedure of Test Method 2, MIL-STD-471, and 13 demonstrations were based on the procedure of Test Method 3, MIL-STD-471. The only other significant applications were the five demonstrations based on Test Method 1 of MIL-STD-471.

Comments on the statistical aspects of the procedures used thus apply generally to MIL-STD-471, Test Methods 1, 2, and 3. General comments on management and administrative aspects of Government maintainability-demonstration standards pertain equally to the older MIL-M26512 and MIL-M23313 standards.

#### 2.4.2 Rejection Experience

Seven rejections were reported by the respondents. In only one case, however, was a retest performed -- a step that is recommended in the maintainability-demonstration-plan provisions of MIL-STD-471. In three cases, the test was extended -- a procedure that would normally result in exceeding the designed test risks. In two cases, the applicable requirement was waived.

Since controlling maintainability and providing information for evaluation purposes are major purposes of demonstration, it is necessary to act on rejections in a positive manner. Correcting maintainability design to eliminate causes of rejection and then retesting would seem to be the most prudent course. A simple test extension or waiver of the requirement avoids the central issue.

#### 2.4.3 Purpose of Maintainability Demonstration

Almost half of the respondents indicated that control of achieved maintainability was the major purpose of maintainability demonstration. One-third believed that the information provided

### QUESTIONNAIRE ON MAINTAINABILITY DEMONSTRATION

Information Concerning the Respondent	
Date	_____
Company or Agency	_____
Division or Department	_____
Name	_____
Title	_____
Address	_____ _____
Nature of Business	_____

Please Return to:

ARINC Research Corporation  
a Subsidiary of Aeronautical Radio, Inc.  
2551 Riva Road  
Annapolis, Maryland 21401  
Attention: H.S. Balaban

EXHIBIT 1

QUESTIONNAIRE MAILED TO MAINTAINABILITY  
MANAGEMENT AND ENGINEERING PERSONNEL



MANAGEMENT ASPECTS OF MAINTAINABILITY (M) DEMONSTRATION (continued)										
Question	Suggested Answers (Mark <input checked="" type="checkbox"/> appropriately)							Other Answers or Elaborations		
4 What are your views on the following M aspects of Standards or contracts with which you are familiar? (Check-off columns are provided for 7 items; please use "Elaboration" column to identify the standards or contracts to which you refer):										
	1	2	3	4	5	6	7			
Definitions of Terms	Satisfactory									
	Unsatisfactory									
Test-Conditions Requirements	Satisfactory									
	Unsatisfactory									
Support-Material Requirements	Satisfactory									
	Unsatisfactory									
Test-Personnel Requirements	Satisfactory									
	Unsatisfactory									
Test-Administration and Test-Reporting Requirements	Satisfactory									
	Unsatisfactory									
5 Where in M-demonstration plans do you think information generated during design and engineering tests should be incorporated?	Parameter specification									
	Distribution analyses									
	Sample selection									
	Bayesian tests									
6 In what areas, if any, have you experienced difficulties in M demonstration?	Differences between test and field environments									
	Intermittent failures									
	Multiple failures									
	Lack of supporting materials									
	Repair times of lengths that are not likely to occur in the field									
	Need for extensive resources (money, time, equipment, etc.)									
7 If you have any additional comments, please use page 8 to write them.										

EXHIBIT 1 (continued)

TECHNICAL ASPECTS OF MAINTAINABILITY (M) DEMONSTRATION		
Question	Suggested Answers (Mark <input checked="" type="checkbox"/> appropriately)	Other Answers or Elaborations
1 What is your general impression of specified <u>M</u> values?	Realistic Unrealistic	
2 What do you think is a good approach to <u>M</u> specification?	Use of historical data or similar equipment Allocation based on a higher level requirement - e.g. availability Use of contractor-proposed <u>M</u> values Use of preliminary <u>M</u> predictions	
3 What parameter or combination of parameters should be specified if the maintenance times are log normally distributed?	Mean Median Percentile Variance	
4 What are your preferences in statistical tests? Please use the check-off boxes to rank the suggested answers in each set.	Set A: <input type="checkbox"/> Fixed sample size <input type="checkbox"/> Multiple samples <input type="checkbox"/> Sequential tests Set B: <input type="checkbox"/> Parametric tests <input type="checkbox"/> Non-parametric tests Set C: <input type="checkbox"/> Classical <input type="checkbox"/> Bayesian <input type="checkbox"/> Decision theory Set D: <input type="checkbox"/> Simulated failures <input type="checkbox"/> Actual failures <input type="checkbox"/> Combination of these	

EXHIBIT 1 (continued)

TECHNICAL ASPECTS OF MAINTAINABILITY (M) DEMONSTRATION (continued)		
Question	Suggested Answers (Mark <input checked="" type="checkbox"/> appropriately)	Other Answers or Elaborations
5 What are your views on the following aspects of the six M-demonstration plans in MIL-STD 471: Plan 1:	<div style="background-color: #cccccc; height: 20px; margin-bottom: 5px;"></div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div>	
Plan 2:	<div style="background-color: #cccccc; height: 20px; margin-bottom: 5px;"></div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div>	
Assumptions	<div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div>	
Specified Parameters	<div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div>	
Decision Criteria	<div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div>	
Other	<div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div>	
Plan 3:	<div style="background-color: #cccccc; height: 20px; margin-bottom: 5px;"></div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div> <div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div>	
Assumptions	<div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div>	
Specified Parameters	<div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div>	
Decision Criteria	<div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div>	
Other	<div style="display: flex; justify-content: space-between;"> <div>Satisfactory</div> <div>Unsatisfactory</div> </div>	

TECHNICAL ASPECTS OF MAINTAINABILITY (M) DEMONSTRATION (continued)	
Question	Suggested Answers (Mark <input checked="" type="checkbox"/> appropriately) →
Question 5 (continued)	
Plan 4:	
Assumptions	Satisfactory Unsatisfactory
Specified Parameters	Satisfactory Unsatisfactory
Decision Criteria	Satisfactory Unsatisfactory
Other	Satisfactory Unsatisfactory
Plan 5:	
Assumptions	Satisfactory Unsatisfactory
Specified Parameters	Satisfactory Unsatisfactory
Decision Criteria	Satisfactory Unsatisfactory
Other	Satisfactory Unsatisfactory
Plan 6:	
Assumptions	Satisfactory Unsatisfactory
Specified Parameters	Satisfactory Unsatisfactory
Decision Criteria	Satisfactory Unsatisfactory
Other	Satisfactory Unsatisfactory

TECHNICAL ASPECTS OF MAINTAINABILITY (M) DEMONSTRATION (continued)			
Question	Suggested Answers (Mark <input checked="" type="checkbox"/> appropriately)	Other Answers or Elaborations	
6 Do you think that sample selection based on stratifying the possible <u>M</u> tasks and sampling in proportion to relative task failure rates is a good procedure? Elaborate, please.	Yes  No		
7 Do you have any suggestions on how the following aspects of <u>M</u> demonstration may be improved? Elaborate, please.			
Specifications	Yes No		
Sampling	Yes No		
Statistical Tests	Yes No		
8 Do you know of any published papers that propose good approaches to <u>A</u> demonstration? List, please.	Yes No		
9 If you have any additional comments, please use page 8 to write them.			

EXHIBIT 1 (continued)

ADDITIONAL COMMENTS ON SANITARIABILITY DEMONSTRATION	
Management Aspects	Technical Aspects

EXHIBIT 1 (concluded)

TABLE I

## SUMMARY OF RESPONSES TO MAINTAINABILITY-DEMONSTRATION QUESTIONNAIRE

<u>Affiliation of Respondent</u>		<u>Maintainability Demonstration Experience</u>	
Industry	29	Involved with one or more	50
Military/Government	10	No involvement	7
		No answer	2
<u>Type of Involvement</u>		<u>Number of Rejections</u> 7	
		<u>Action Following Rejection</u>	
Designer of Equipment	14	Test Extended	3
Designer of Maintainability-		Equipment Redesign	2
Demonstration Test	21	Equipment Redesign and Retested	1
Conductor of Maintainability-		Applicable Requirement Waived	2
Demonstration Test	19	Penalty Provision Invoked	0
Monitor of Maintainability-		Other	1
Demonstration Test	16		
<u>Standards Used</u>			
MIL-STD-471		MIL-STD-471	
Plan 1	5	Plan 4	1
Plan 2	9	Plan 5	0
Plan 3	4	Plan 6	1
		MIL-M26512	8
		MIL-M23313	9
		Other	4
<u>Purpose of Maintainability Demonstration</u>			
Provide accept/reject criteria			7
Provide maintainability information			13
Provide control on achieved maintainability			18
Verify maintainability prediction			8
Other			2
<u>Views on Elements of Maintainability Standards or Contracts</u>			
<u>Element</u>			<u>Satisfied</u> <u>Unsatisfied</u>
Definition of Terms		39	37
Test Condition Requirements		37	30
Support Material Requirements		44	22
Test Personnel Requirements		35	28
Test Administration and Reporting Requirements		45	23
<u>Use of Prior Information</u>			
Parameter Specification		12	
Distribution Analyses		7	
Sample Selection		24	
Bayesian Tests		4	
Other		2	
No Response		5	

(continued)

TABLE I (continued)

<u>Difficulties Experienced</u>			
Differences between test and field environments			19
Intermittent failures			9
Multiple failures			4
Lack of supporting materials			7
Abnormally large repair times			6
Extensive resources required			16
Other			2
No Response			6

<u>Impression on Specified Maintainability Values</u>			
	<u>Military</u>	<u>Industry</u>	<u>Total</u>
Realistic	4	17	21
Unrealistic	5	10	15
No Response	2	2	4

<u>Approach to Maintainability Specification</u>	
Historical Data	20
Allocation	28
Contractor Values	5
Preliminary Maintainability Predictions	9
Other	2

<u>Lognormal-Parameter Specification</u>			
Mean	4	Median and Percentile	4
Median	2	Median and Variance	3
Percentile	1	Mean, Median, Percentile	2
Mean and Median	2	Mean, Median, Variance	1
Mean and Percentile	7		
Mean and Variance	10		
Total Mean		26	
Total Median		14	
Total Percentile		14	
Total Variance		14	

<u>Type of Statistical Test</u>					
<u>Test</u>	<u>Rank 1</u>	<u>Rank 2</u>	<u>Rank 3</u>	<u>Average Rank</u>	<u>Preference</u>
A. Fixed Sample	13	13	9	1.89	2
Multiple Sample	6	22	7	2.03	3
Sequential	13	14	4	1.63	1
B. Parametric	22	11	—	1.26	1
Nonparametric	11	22	—	1.67	2
C. Classical	16	6	5	1.59	1
Bayesian	3	17	7	2.15	3
Decision Theory	8	14	5	1.89	2
D. Simulated Failures	15	17	5	1.73	2
Actual Failures	8	14	15	2.19	3
Combination	14	20	3	1.70	1

(continued)

TABLE I (concluded)

Views on MIL-STD-471								
Plan Number	Rating [Satisfactory (S) or Unsatisfactory (U)]							
	Assumptions		Specified Parameters		Decision Criteria		Other	
	S	U	S	U	S	U	S	U
1	23	4	21	7	21	7	7	6
2	22	5	23	5	19	8	7	5
3	17	7	19	6	18	4	7	4
4	19	0	16	4	17	1	7	5
5	14	2	13	4	13	2	7	3
6	16	3	17	3	17	2	6	2
<p align="center"><u>Approval of Stratification</u></p> <p align="center">Yes                      17</p> <p align="center">No                         10</p>								
<p><b>NOTE:</b> Many questionnaires contained detailed comments on one or more questions; these are not included in this tabulation.</p>								

is most important. These two answers, of course, are not mutually exclusive; i.e., it can be assumed that the information can and will be used to control maintainability.

Eight respondents believed that maintainability demonstration is most useful for verifying a maintainability prediction. In one sense, this would imply that the prediction is more pertinent to the maintainability program than actual test results. This would be the case, for example, if the field maintenance and logistic planning depended heavily on the maintainability prediction.

In practice, however, the maintainability demonstration is geared to the contractual maintainability requirements, and the prediction is another element of a maintainability program to provide control. If a prediction indicated that the contractual requirement would not be met, it would be unwise to proceed to the demonstration without carefully viewing the requirement, the design, and the prediction procedure and inputs to make any necessary revisions.

#### 2.4.4 Views on Aspects of Maintainability Standards or Contracts

As can be seen in Table I, the percentage of respondents who were satisfied with the maintainability aspects listed in the questionnaire is not encouraging. Almost half of the responses indicated that one or more contracts or standards had unsatisfactory definitions. Test-condition requirements and test-personnel requirements also received relatively unsatisfactory ratings. While support material and test-administration and -reporting requirements were rated better than the others, about one-third of the ratings were unsatisfactory.

These results vividly illustrate the need for careful attention to the specifics of definitions and requirements concerning maintainability demonstration in contractual documents and in associated military standards.

#### 2.4.5 Use of Prior Information

As would be expected, the two most prominent categories given for the use of prior information in maintainability demonstration were sample selection and parameter specification. Current sampling procedures are usually a form of proportional stratified sampling, which requires identification of tasks, means for grouping tasks, and relative frequency of occurrence of the groups. Parameter specification may require information on (1) a higher-level requirement such as availability, (2) the type of maintenance-time distribution, and (3) achieved maintainability on similar systems or equipments.

#### 2.4.6 Difficulties Experienced

The two difficulties mentioned most often in the survey were differences between the test environment and the field environment, and the extensive resources required for planning and conducting the demonstration.

Differences between test and field environments will always be present, but every effort should be made to minimize them through careful planning. On the other hand, there may be cases in which duplication of environment will be costly and difficult.

A compromise approach is to know what the environmental differences are and to adjust the specified parameter values accordingly.

If the maintainability demonstration is considered merely an exercise in statistics, with little or no enforcement, the expenditure of funds is essentially wasteful. If the contractor knows that failure to pass the demonstration test will require further efforts on his part to improve the maintainability design, he will have the incentive to provide carefully for the maintainability demonstration.

Economy, of course, is still a major objective; careful planning and integration with the complete test program is essential. Sample sizes can be reduced by using the most efficient tests or by accepting higher risks. Use of prior information in a Bayesian test is one new approach to limiting the amount of necessary testing.

#### 2.4.7 Opinions on Specified Maintainability Values

Almost 50 percent of the respondents believed that specified maintainability values were realistic. It was noted, however, that military and Government personnel were less convinced of this realism (45 percent) than industry personnel (63 percent). Generally, it should be the customer (i.e., military) who specifies quantitative values, but in practice the contractor may play a prominent role in such specifications.

#### 2.4.8 Approach to Maintainability Specifications

Four alternative approaches to specifying maintainability values were presented, and approximately 44 percent of the respondents favored an approach based on an allocation of a higher-level requirement such as availability or effectiveness. Since most of the respondents were from industry, this might be a reflection of their desire for flexibility, which might be stated as follows: "Tell us your overall objective; let us worry about how we will achieve it".

In many cases, this might be a prudent policy, but there are situations in which complete freedom is undesirable. For example, the following sets of MTBF and MTTR values will both yield the same steady-state availability of 0.90:

• Set 1: MTBF 9; MTTR 1

• Set 2: MTBF 90; MTTR 10

The cost of the two sets, however, can differ greatly. The first set requires, on the average, 10 times the number of maintenance actions as the second for continuously operated systems; and if maintenance consists generally of replacement, Set 1 will require much greater inventory levels. On the other hand, if mission time is small, for example, less than one hour, a 9-hour MTBF may be acceptable, but a 10-hour downtime may be operationally unacceptable from the viewpoint of readiness. If the system is not operated between missions, Set 1 may be preferable.

Examples such as these emphasize the need for as complete a specification as possible for the overall requirement and associated constraints. The application of an allocation procedure under these conditions is generally a good policy, because it provides assurance that requirements will be consistent with the overall objective.

The second most frequently mentioned approach to maintainability specification was the use of historical data. This is a natural means for imposing realistic requirements. It is generally unwise to specify a maintainability value without reference to the state of the art.

Of course, for new equipments, maintenance philosophies, and operational procedures, there may be little relevance to past history, and it is in such cases that problems arise. In these situations, conservatism may be the most prudent course, because the risk of failure is too great if large improvements over the state of the art are expected.

It may also be possible to provide some flexibility in the requirement so that as development progresses and exploratory-test results are evaluated, a realistic maintainability requirement for demonstration can then be invoked upon agreement between contractor and customer. Such a policy will require some form of restriction in the initial contract (e.g., minimum and maximum values) to protect both parties.

A third approach for new types of systems and maintainability policies is an incentive arrangement whereby only a minimum maintainability value is specified. The demonstration

procedure may then be an estimation process in which the incentive payment or penalty is geared to the estimate (either point or interval).

#### 2.4.9 Lognormal-Parameter Specification

To the question, which parameter or parameter combination should be specified if maintenance times were distributed log-normally, the most frequent answer (10 of 36 responses) was the mean/variance combination. This is surprising because (1) no currently used method is based on this combination; and (2) the variance of a lognormal distribution, unlike the median and percentile values, is difficult to interpret.

Of the two central-tendency parameters, the mean was included in 26 responses and the median in 14. This is in agreement with the response to the previous question, in which allocation was the preferred approach for specification, since the mean possesses more desirable properties than the median for allocation.

The mean/percentile combination was the next preferred specification, and this combination corresponds to several MIL-STD-471 plans.

Three responses referred to three lognormal parameters, but since any two of the three parameters listed dictate the value of the others, the three-parameter specifications have little justification.

#### 2.4.10 Type of Statistical Test

From the four categories considered, the most preferred plan (assuming the responses were independent for each category) would be a classical, sequential nonparametric test based on a combination of simulated and naturally occurring failures. No currently used test procedure is a nonparametric sequential procedure, although Test Method 1 can be considered such a test since the lognormal assumption is used primarily for converting the lognormal-parameter specification to a binomial-parameter specification.

The nonparametric and classical-type tests are the clear choices for the alternatives presented. The sequential test is only slightly preferred over the fixed-sample-size test, while the use of simulated failures only is almost as preferred as a combination of simulated and actual failures.

#### 2.4.11 Views on MIL-STD-471

All plans received relatively good ratings on the three aspects listed. There was surprisingly little comment on the risk

errors in the MIL-STD-471 methods. Many comments were related to (1) the inflexibility of some of the methods, e.g., a fixed sample size for methods 3 and 4; (2) inappropriateness of the lognormal assumption for specific types of equipments; and (3) the need for large sample sizes.

The comment about inflexibility is valid since the MIL-STD-471 plans do not generally provide for varying risks. The log-normal assumption may not hold for a specific type of system or maintenance comment, and that is why a nonparametric test is included in the standard. However, other nonparametric tests are available that would offer more flexibility.

The sample size (e.g., 50 for Test Method 4) is based solely on the risks. The only way to reduce the sample size for a given procedure is to increase risks (or effectively have less discrimination in the test between acceptable and unacceptable maintainability). The use of more efficient test procedures -- sequential rather than fixed, parametric rather than nonparametric, and Bayesian rather than classical tests -- is another alternative that can lead to reduced sample sizes.

#### 2.4.12 Approval of Stratified Sampling

Twenty-seven of 37 responses indicated approval of stratified sampling procedures. There were, however, many comments on stratification, with many different suggestions for improving current task-selection procedures. Very little comment was made on the current use of analytical procedures based on simple random sampling when, in fact, the sample observations are usually obtained through a stratified procedure.

The reason most often given for favoring stratified sampling was that it is the best way to ensure representativeness.

### SECTION III

#### DEMONSTRATION AS A MAINTAINABILITY-PROGRAM ELEMENT

##### 3.1 GENERAL

In Section I, the requirement for demonstration as a maintainability-program element as specified in MIL-STD-470 was noted. The major points of paragraph 5.11 of MIL-STD-470 are as follows:

- (1) Maintainability demonstration is a contractual requirement.
- (2) Maintainability demonstration will normally be performed in accordance with MIL-STD-471.
- (3) The contractor will submit a maintainability-demonstration plan.
- (4) The maintainability-demonstration test will be integrated with other system requirements.
- (5) The formal demonstration will be conducted in an operational or simulated environment.

With regard to item 2, MIL-STD-471 contains the detailed procedure, but of particular pertinence to this section are the major elements to be included in the maintainability-demonstration plan, which are listed in Table II.

TABLE II

#### MAJOR ELEMENTS OF A MAINTAINABILITY- DEMONSTRATION PLAN (MIL-STD-471)

Description of Demonstration Conditions
Description of Test Team
Description of Demonstration Support Material
Predemonstration-Phase Schedule
Description of Formal Demonstration Test
Retest-Phase Schedule

Timely planning and careful management of the maintainability demonstration is required to fulfill the requirements of MIL-STD-470 and 471. This chapter presents a general discussion of some of the major problems that can occur in such planning and management and offers guidelines for selecting and implementing the best policies.

### 3.2 MANAGEMENT RESPONSIBILITIES IN MAINTAINABILITY TESTING

#### 3.2.1 Development-Test and Demonstration-Test Management

The maintainability-demonstration procedure is generally the last of a series of tests and evaluations concerned with maintainability parameters. Each of these tests and evaluations is designed to provide information about system maintainability. Tests designed to aid in improving or achieving desirable maintenance characteristics can be classified as information tests or development tests. The management aspects of development tests and maintainability-demonstration tests are summarized in Tables III and IV, respectively.

#### 3.2.2. Overall Test Planning

It is the responsibility of management to plan and manage the necessary tests in such a manner that pertinent and timely information concerning maintainability achievement will be provided as economically as possible. This is not an easy task in view of the many unknowns that exist in a typical system-development program.

An overall test plan should be developed at the beginning of the program. At this point, the plan will not be completely detailed. However, it should provide for determining specific test requirements and procedures at some point before the testing is to be performed, allowing the review of all aspects of the tests by interested design, engineering, and management groups.

Since the features of many types of tests, particularly those in the development stage, will depend on the achievements made and problems encountered during the design-evaluation process, the test program must be flexible enough to allow appropriate changes in test parameters, procedures, and decision rules. Such changes, however, should be carefully controlled and reviewed by the cognizant Air Force agency. Table V summarizes the steps necessary for overall test planning.

### 3.3 THE MAINTAINABILITY-DEMONSTRATION-TEST PLAN

The decision to require a formal maintainability-demonstration test rests with the Air Force and should be based on tactical considerations, mission requirements, cost of tests,

TABLE III  
MANAGEMENT ASPECTS OF DEVELOPMENT TESTS

Purpose of Tests	To determine physical realizability, to determine functional capabilities, to establish the basic design.
Responsible Groups	Air Force or contractor's design-engineering group, with support from other groups as required.
General Description	Development tests are usually informal exploratory tests designed to provide fundamental R&D information about a basic design. Nominal environmental levels are used unless the test is specifically oriented to check for effects at environmental extremes. Sample sizes are limited, but the general principles of good experimental and statistical design should be followed.
Examples of Specific Types of Tests	<ul style="list-style-type: none"> <li>• Component Accessibility</li> <li>• Fault-Isolation Routines</li> <li>• Test-Point Adequacy</li> <li>• Harmonization Requirements</li> <li>• Compatibility Tests</li> </ul>
Test Scheduling	Not usually specified formally. Design-engineering group establishes schedules to meet design-development objectives. Such schedules must conform to development-program milestones.
Test Items	Basic materials, off-the-shelf parts and assemblies, prototype hardware.
Test Documentation	Engineering test reports and analyses. Maintainability information to be documented for later use in prediction, evaluation, and testing tasks.
Test Follow-Up Action	Determination of design feasibility or need for redesign. Implementation of test information in further design work. Approval, modification, or disapproval of design, materials, and parts.
Maintainability Provisions	Proposed materials and designs to yield acceptable maintainability performance are tested on limited samples. Packaging tests and component-interaction tests are examples. All maintainability data should be fully documented for future use in prediction, assessment, and later testing activities.

**TABLE IV**  
**MANAGEMENT ASPECTS OF MAINTAINABILITY-DEMONSTRATION TESTS**

**Purpose of Tests**

To demonstrate formally that the maintainability requirements are achieved.

**Responsible Groups**

Air Force or contractor's program-management, maintainability, and effectiveness-assurance groups, with tests monitored and approved by the Air Force.

**General Description**

Demonstration tests are performed on the major end items, often at the highest system level, under realistic operational and environmental conditions. Rules are specified for classifying failures, performing repairs, allowing design changes, etc. Time is an inherent test parameter. The test design is usually directed towards providing a specified statistical confidence for making an appropriate decision.

**Test Scheduling**

Demonstration-test schedules are normally contract-specified. They generally occur before full-scale production but after initial production, when test samples are available.

**Test Items**

Production hardware at major end-item level.

**Test Documentation**

Contract-specified procedures or clause requiring contractor to submit complete test plan. Test results fully documented, including analyses and conclusions concerning the meeting of contract requirements.

**Test Follow-Up Actions**

Acceptance or rejection of equipment with respect to maintainability requirements. Failure to pass demonstration tests will require appropriate design and assurance efforts on the part of the contractor.

**Maintainability Provisions**

Demonstration tests are specifically designed to test for maintainability and associated parameters at the equipment and system levels.

**TABLE V**  
**STEPS IN OVERALL TEST PLANNING**

1. Determine test requirements and objectives
2. Review existing data to determine if any existing requirements can be met without tests
3. Review a preliminary list of planned tests to determine whether economies can be realized by combining individual test requirements
4. Determine the necessary tests
5. Allocate time, funds, and effort to perform these tests
6. Develop test specifications at an appropriate level, or make reference to applicable sections of the system specification to provide direction for later development of test specifications
7. Assign responsibility for test conduct, monitoring, analysis, and integration
8. Develop review and approval policies for test-reporting procedures and forms
9. Develop procedures for maintaining test-status information throughout the entire program

and an evaluation of the likelihood of achieving the maintainability requirement simply as a result of good design procedures without a demonstration test. It is essential to recognize that a maintainability-demonstration test does not guarantee achieving the required maintainability. It focuses the contractor's attention on maintainability, but often this is not sufficient unless penalties for test failure are included in the contract.

### 3.3.1 Responsibility for Preparing Test Plan

There are no fixed rules for determining Air Force and contractor responsibilities in preparing maintainability-demonstration-test plans. The peculiarities of individual procurements require flexibility in the requirements for supplying information. MIL-STD-471 outlines certain requirements for a maintainability-test plan. Generally the plan is submitted as

part of the contractor's proposal and includes details that are applicable to maintainability information supplied by the Air Force. The specific information supplied to the contractor by the Air Force should meet both of the following criteria:

- It should provide direction for developing the test design.
- It should be based on operational or tactical constraints or on a trade-off analysis.

In most cases, these criteria will be met if the following are provided:

- Equipment configuration for test
- Maintenance concept
- Maintenance environment
- Levels of maintenance to be demonstrated
- Modes of operation for the test
- Test team organization

### 3.3.2 Contractor's Responsibility

In his proposal, the contractor must include certain information about the maintainability-demonstration-test plan, regardless of whether the information has been derived from the Air Force or the contractor. The plan submitted by the contractor to the Air Force should include the elements listed in Table VI, which was extracted from MIL-STD-471.

### 3.3.3 Maintainability-Demonstration-Plan Milestones

The major milestones in the development of a final maintainability-demonstration plan are listed in Table VII, in which it can be seen that the maintainability-demonstration plan must be continually updated as system development progresses, to reflect changes in requirements and design and to incorporate the results of the maintainability-program design reviews, predictions, and assessments.

Specific dates for review by the procuring activity should be established at the time the contract is awarded. Before the test is conducted, the maintainability-demonstration plan and detailed procedures must receive final approval of the reviewing activity.

**TABLE VI**

**ELEMENTS OF CONTRACTOR'S MAINTAINABILITY-DEMONSTRATION TEST**

**Description of Demonstration Conditions**

- Quantitative maintainability requirements
- Maintenance concept
- Maintainability-demonstration environment
- Levels of maintenance to be demonstrated
- Demonstration sites and facility requirements
- Participating agencies
- Mode of operation for the test
- Item(s) to be demonstrated

**Description of Test Team**

- Organization
- Degree of participation for contractor and procuring activity
- Assignment of specific responsibilities
- Qualification, quantity, and training of test-team personnel

**Description of Demonstration Support Material**

- Support equipment
- Tools and test equipment
- Technical publications
- Spares and consumables
- Safety equipment
- Calibration support requirements

**Predemonstration-Phase Schedule**

- Assembly of test team
- Training
- Preparation of facilities and support material

**Description of Formal Demonstration Test**

- Test objectives
- Schedule of tests
- Task-selection method
- Test method
- Data-acquisition method
- Analytical and calculation methods
- Specific data elements
- Time units of measurement
- Type and schedule of reports
- Description and schedule of preventive-maintenance tasks
- Description of corrective-maintenance tasks

**Retest Phase**

**TABLE VII**  
**MILESTONES IN THE DEVELOPMENT OF THE**  
**MAINTAINABILITY-DEMONSTRATION PLAN AND PROCEDURE**

Period	Input	Output for Maintainability-Demonstration Applications
Preproposal and Proposal	Initial work statement	Proposed maintainability-demonstration plan
Contract Award	Final technical and cost requirements	Updating of maintainability-demonstration plan
System Development	Maintainability design; development-test results; maintainability predictions and analyses	Revisions to maintainability demonstration plan
Final Test Planning	Overall system-test program	Final maintainability-demonstration plan; integration with overall test program
Predemonstration Phase	Maintainability-demonstration plan	Test team, facilities, and support material assembled; detailed procedural methods for sampling, analyzing, and reporting results completed
Maintainability Demonstration	Approved maintainability-demonstration plan and procedures; equipment, facilities, material and personnel	Results of maintainability-demonstration tests
Retest Phase (if necessary)	Results of maintainability-demonstration tests; retest plan	Results of retest

### **3.4 RELIABILITY AND MAINTAINABILITY IMPROVEMENT THROUGH DEMONSTRATION**

Although the basic purpose of maintainability demonstration is control, ARINC Research Corporation has found, during the course of several demonstration-monitoring studies, that an important by-product is the discovery of faults and procedures that degrade an equipment's reliability and maintainability.

The comments in Table VIII were extracted from several ARINC Research letter reports to military customers presenting the observations of ARINC Research personnel monitoring maintainability demonstrations. It can be seen from these comments that the maintainability demonstration can provide information leading to improvements in manuals and support equipment, specific circuit design, overall maintenance-design philosophy, packaging, and sparring. Because of this desirable by-product of the maintainability-demonstration effort, the customer should plan to have knowledgeable representatives monitoring the test to discover design deficiencies and recommend means for improvement.

Contractual allowance for remedying serious deficiencies should also be considered. A specific maintenance design or procedure may not cause excessive downtime but may make the equipment unsafe or highly failure-prone during operational use. It is thus important to recognize that the passing of the demonstration test does not necessarily mean that major or critical improvements cannot or should not be made.

Of course, if the equipment passes the test, the contractor may not be obliged to act on recommendations resulting from the demonstration, but may do so for changes that can be easily incorporated. If the recommended changes will yield a significant improvement, contract renegotiation may be called for. In any event, for additional procurements, such as full-scale production, the customer should act positively on the information provided by the results of the demonstration-test monitoring.

## TABLE VIII

### COMMENTS OF MAINTAINABILITY-DEMONSTRATION OBSERVERS

1. "No special support equipment was used during the demonstration to isolate the inserted faults. A module puller was used, however, to remove the modules. This item is not required, but assists in module removal and minimizes possible module damage. It is recommended that a module puller be procured and made available to the customer."
2. "It is recommended that future procurements address a requirement for greater ease in the removal and replacement of terminal assemblies than now exists. Assembly guide rails and guide pins, securing devices requiring no special tools, and use of receptacles and plugs in lieu of terminal boards and taper pins, where appropriate, are cases in point. A form factor permitting convenient removal of assemblies from the front of the mounting racks (without disturbing other assemblies) is an important feature for field installations."
3. "The technicians relied almost exclusively on the handbook signal-flow diagrams for fault isolation during the demonstration, and they commented that the diagrams were superior to conventional schematic diagrams. On the basis of their effectiveness in the demonstration, it is recommended that signal-flow diagrams be incorporated extensively in all organizational-level maintenance publications."
4. "Individual maintenance modules of the AN/XYZ are susceptible to damage when removed from the equipment. It is recommended that suitable containers be developed to protect modules in transit to and from repair facilities."
5. "Replacement of the Tube/Shield Assembly (AB123) is difficult and time-consuming. During the demonstration, replacement time exceeded 45 minutes for two technicians working together.  
  
"It is recommended that captive hardware be used to attach the faceplate and the Tube Shield Assembly and that the yoke and ground wires be connected to chassis wiring by a bayonet-type connector."
6. "Connectors C3, C5, and C6 are all the same size, are located side-by-side, and have identical external-plug housing keying. However, the internal pins of each connector are arranged in patterns different from other connectors

(continued)

TABLE VIII (continued)

so that only the correct plug will fit the connector. A problem arises in that the plug housings will fit the keying on any of the four connectors, and only the connector-pin pattern prevents the plug mating. If sufficient force is applied, pins can be damaged.

"It is recommended that the keying on all four of these plugs and connectors be changed so that each pair is unique, preventing the possibility of pin damage."

7. "The module guides in the AN/FJK equipment used in the maintainability demonstration (serial number 1) were very weak. One guide was broken off and allowed a module to be inserted skewed. The module was improperly seated, which resulted in a diode failure. Other module guides were loose and allowed the modules to be improperly seated in their plug-in connectors. The contractor has noted this design deficiency, and has incorporated an improved module guide that is much stronger than the guides used in the serial number 1 equipment. The contractor stated that all subsequent production units will contain the new improved module guides."
8. "The front-access modules (M1, M2, M3) are retained in place by the hinged face plate of the indicator. Technicians on two occasions during the demonstration failed to fully insert the Selection/Deflection Amplifier (M3) resulting in improper equipment operation. In all cases, closing of the hinged front panel was difficult because of interference between the Tube/Shield Assembly (M1) and the bezel gasket assembly.  
  
"It is recommended that front-access modules be retained by fasteners (preferably quick-release captive type) to ensure full insertion of modules in their receptacles, and that the bezel/module interference be corrected."
9. "Lamp Driver Assembly (L6) circuit-card replacement is extremely difficult. The technician must lean over the extended Display Tube Subassembly and, at the limit of his reach, release two quarter-turn fasteners with a screwdriver. He must then slide the assembly toward himself and rotate it approximately 135 degrees on its cables for access to the circuit cards. The circuit cards require excessive force for removal and insertion. During the demonstration, a technician broke one of the cable wires while replacing a circuit card even though he exercised great care.

"It is recommended that the Lamp Driver Assembly be repackaged for satisfactory in-place maintenance."

## SECTION IV

### THE STATISTICAL BASIS OF MAINTAINABILITY-DEMONSTRATION

#### 4.1 GENERAL

A maintainability-demonstration test provides the information necessary for accepting or rejecting the product with respect to conformance to stated maintainability goals. The demonstration effort, therefore, provides the input to a decision process. Since this input is statistical in nature -- that is, the results of the test are not constant, but are subject to statistical (random) fluctuation -- it is necessary to consider how such fluctuation can result in incorrect decisions.

In practice, an accept or reject decision may not be the only alternative (e.g., the decision may be to "accept" after a specified design change is made). However, it is easiest to look at the maintainability-demonstration effort from the accept/reject viewpoint to avoid having to consider all possible types of decisions that can be made.

#### 4.2 DECISION ERRORS

If the discussion is restricted to accept/reject decisions, two basic types of errors can be made:

Type I Error: Reject the equipment for not meeting its maintainability requirement when, in fact, it has.

Type II Error: Accept the equipment as meeting its maintainability requirement when, in fact, it has not.

The Type I error is detrimental primarily to the producer since his acceptable equipment is being rejected. The consumer also wants to minimize this error since he often has immediate need for the equipment and would not want to experience the unnecessary delay caused by this type of error. In the terminology of acceptance testing, this error is often denoted by alpha ( $\alpha$ ) and is called the producer's risk.

The Type II error is detrimental primarily to the consumer since he is accepting equipment that is below standard. Unless there is a contractual requirement for the producer to maintain acceptable operational performance, he has no direct interest in the Type II error. This type of error is called the consumer's risk and is usually denoted by beta ( $\beta$ ).

It is emphasized that for any type of decision procedure in which acceptable and unacceptable levels of maintainability can be defined, the Type I and Type II errors are always in force. The statistical theory of acceptance testing allows for controlling such errors by controlling the sample size, selecting appropriate test statistics, and invoking an applicable decision criterion.

#### 4.3 THE TEST HYPOTHESIS

Because of the statistical nature of demonstration testing, the basis for decision is developed from the statistical theory of hypothesis testing. In maintainability testing, the hypothesis under test -- called the null hypothesis and denoted by  $H_0$  -- is usually that the submitted product conforms to the maintainability requirement. An alternative hypothesis, denoted by  $H_1$ , is also specified (or implied); it states that the product is at an undesirable maintainability level. Rejection of the null hypothesis is equivalent to acceptance of the alternative hypothesis.

For example, assume that the parameter of interest is mean corrective-maintenance time, denoted by  $\bar{M}_{ct}$ . Then two basic forms of test hypotheses are as follows:

##### TEST I

$$H_0: \bar{M}_{ct} = 30 \text{ min}$$

$$H_1: \bar{M}_{ct} = 60 \text{ min}$$

##### TEST II

$$H_0: \bar{M}_{ct} = 30 \text{ min}$$

$$H_1: \bar{M}_{ct} > 30 \text{ min}$$

For Test I,  $H_0$  and  $H_1$  specify unique values for  $\bar{M}_{ct}$  and with respect to this characteristic are called simple. Strictly speaking, a simple hypothesis is one that completely specifies the distribution of the random variable. If maintenance time is exponential -- i. e.,  $f(x) = \frac{1}{\theta} e^{-x/\theta}$  --  $\theta$  is uniquely determined by a mean specification such as those of Test I. For a two-parameter distribution such as the normal or lognormal, neither  $H_0$  nor  $H_1$  of Tests I and II uniquely determines the distribution. For Test II,  $H_1$  does not specify a unique value and is, therefore, called composite with respect to  $\bar{M}_{ct}$ .

The relationship between the test hypotheses and the decision errors is now evident. For both Tests I and II above,  $H_0$  represents the more desirable maintainability level, and

the rejection of  $H_0$  when it is true would represent the producer's risk  $\alpha$ .  $H_1$  for Test I represents an undesirable level of maintainability, and acceptance of an equipment (equivalent to accepting  $H_0$ ) when  $H_1$  is true is a risk to the consumer. Test II, however, does not specify a unique value for  $H_1$ ; therefore, the consumer's risk cannot be evaluated.

If the Test I hypotheses were to be invoked and the necessary distribution assumptions on maintenance time were validated, the assignment of  $\alpha$  and  $\beta$  risks corresponding, respectively, to rejecting  $H_0$  when it is true and accepting  $H_0$  if  $H_1$  is true would generally be sufficient for determining the sample size,  $n$ , test statistic, and decision criterion to meet these risk levels.

If the Test II hypotheses are to be invoked, the general procedure is to use as large a value of  $n$  as possible, since this will minimize the  $\alpha$  risk as well as minimizing the chance of accepting  $H_0$  when it is false. This conclusion is simply a result of the fact that the variance of an estimate generally decreases as the sample size increases, and, thus the larger  $n$  is, the more precision in a sample statistic and the less risk of an incorrect decision.

Because of the importance of the test hypotheses and associated risks, a separate section (Section V, The Maintainability Demonstration Specification), is presented to provide guidelines for parameter and risk specification.

#### 4.3.1 Specification in Terms of Confidence Level

Test requirements are sometimes specified in terms of confidence levels. Such a specification, however, is subject to serious misinterpretation, as the following example illustrates. Assume that the specification states: "...a sample shall be tested to determine with 90-percent confidence that the equipment conforms to the requirement of a 1/2-hour mean time to perform corrective maintenance. . ." Two reasonable test criteria are:

Test I: Compute the 90-percent lower confidence limit,  $\mu_L$ . Since there is 90-percent confidence that the true mean time to repair is greater than  $\mu_L$ , if  $\mu_L > 1/2$ , reject the equipment; otherwise, accept it.

Test II: Compute the 90-percent upper confidence limit,  $\mu_U$ .

Since there is 90-percent confidence that the true mean time to repair is less than  $\mu_U$ , then if  $\mu_U \leq 1/2$ , accept the equipment; otherwise, reject it.

Test I is equivalent to one in which the producer's risk is 10 percent at a true mean of  $1/2$  hour. Test II is equivalent to one in which the consumer's risk is 10 percent at a true mean of  $1/2$ -hour. The difference between the two tests is apparent: The former requires that equipment with a mean corrective-maintenance time of  $1/2$ -hour be accepted 90 percent of the time; the latter requires acceptance only 10 percent of the time if the equipment  $M_{ct}$  is  $1/2$ -hour.

Most specifications of this form are designed to represent a Test II criterion. This criterion makes no provision for the producer's risk at a highly acceptable maintainability level. Many plans will meet the criteria of Test II. Generally, the lower the sample size the higher the producer's risk for a fixed confidence and maintainability level.

In any case, if a test specification is to be made in the form of a confidence interval, it is imperative that it be made clear whether the maintainability numeric represents an acceptable or unacceptable maintainability level.

#### 4.4 RELATIONSHIP OF RISKS TO SAMPLE SIZE

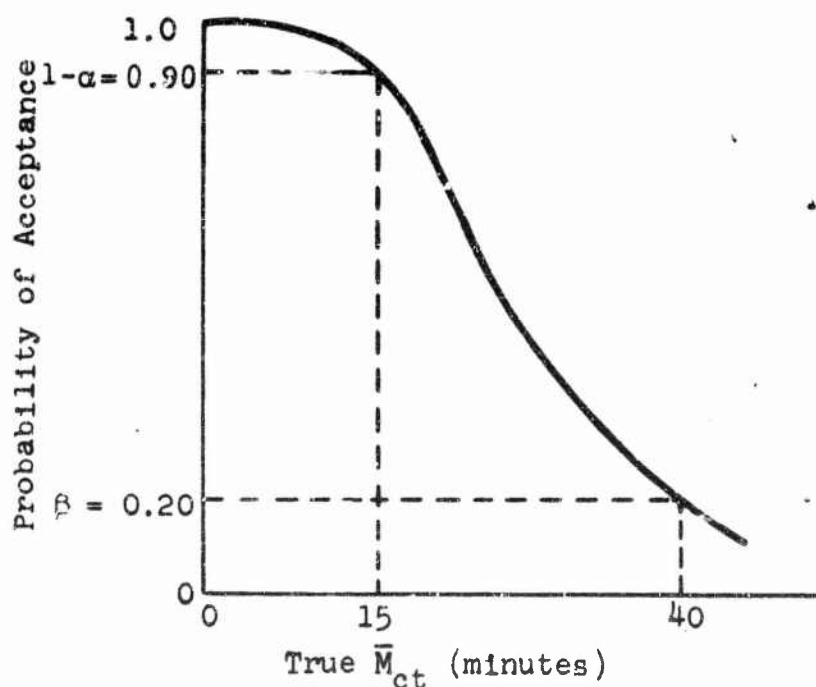
For most tests, the magnitude of  $\alpha$  and  $\beta$  and the number of test observations,  $n$ , are interrelated in such a manner that specifying any two of the quantities determines the third. In the past, for nonsequential tests,  $\alpha$  and  $n$  were usually specified and a test was chosen to minimize the  $\beta$  error. For acceptance testing, the trend now is to specify  $\beta$  instead of  $\alpha$ . If it is important that both  $\alpha$  and  $\beta$  be specified, the sample size is no longer at the discretion of the experimenter, as shown for many of the fixed-sample-size plans presented in Section VII. In sequential sampling,  $\alpha$  and  $\beta$  must be specified in advance, and the sample size is a random variable since its value is not predetermined but will vary over successive tests.

#### 4.5 THE OPERATING-CHARACTERISTIC (O.C.) CURVE

By specifying two of the three quantities  $n$ ,  $\alpha$ , and  $\beta$ , the accept-reject criterion of the acceptance test is uniquely determined for a given family of tests (e.g., a fixed sample test under the lognormal assumption with known variance). It is then possible to generate the O.C. curve of the test plan. This curve shows the probability of acceptance over all possible incoming maintainability levels. Two points on the O.C. curve are already determined -- the  $\alpha$  and  $\beta$  points with their corresponding maintainability levels, which are given by  $H_0$  and  $H_1$ , respectively.

For example, assume that the specification is in terms of  $\bar{M}_{ct}$ , the mean corrective maintenance time, and that  $H_0$  is  $\bar{M}_{ct} = 15$  minutes and  $H_1$  is  $\bar{M}_{ct} = 40$  minutes. The  $\alpha$  risk is 0.10, and the  $\beta$  risk is 0.20. The general shape of the O.C. curve will then be as shown in Figure 1.

The probability of acceptance can be interpreted as the long-run proportion of equipment or lots that will be accepted. If, for example, the O.C. curve shows that an  $\bar{M}_{ct}$  of 25 minutes will be accepted with a probability of 0.65, then in the long run 65 percent of all incoming products with a 25-minutes  $\bar{M}_{ct}$  will be accepted.



For the acceptance test:  $H_0: \bar{M}_{ct} = 15$  minutes  
 $H_1: \bar{M}_{ct} = 40$  minutes,  $\alpha = 0.10$ ,  $\beta = 0.20$ .

FIGURE 1  
 TYPICAL O.C. CURVE

## 4.6 TYPES OF MAINTAINABILITY DECISION TESTS

A decision test may be classified in many different ways. Some of the more important categorizations of maintainability tests are discussed below.

### 4.6.1 Type of Measurement

Measurement type is usually categorized in terms of attributes or classification data and variable or measurement data.

In maintainability testing, the usual attribute-type test is one in which a success/failure determination is made on each sample observation according to some pre-established criterion. Thus if the maintainability requirement is related to maximum duration of repair times, a corresponding attribute measurement is that a particular maintenance-task observation did or did not exceed a specified maximum time. The actual time spent on the task is not directly used in the decision criterion.

A variable measurement, on the other hand, does employ actual measurement of a random variable that is continuously distributed. For maintainability tests, such random variables are usually maintenance times.

For cases in which either type of measurement may be employed, such factors as type of information provided, degree of protection afforded, amount and cost of inspection, and ease of administration should be considered. Table IX summarizes the advantages and disadvantages of each type of measurement with respect to these considerations.

### 4.6.2 Type of Maintenance-Task-Sample Selection

Testing a system's maintainability requires sampling from the various possible types of maintenance tasks that comprise the hypothetical total population of maintenance tasks. The major alternatives to consider are whether induced (simulated) failures or naturally occurring failures are to be considered and, if the former method is chosen, whether simple random sampling is appropriate. Section VI covers these alternatives in detail.

### 4.6.3 Single, Multiple, and Sequential Sampling

Single, multiple, and sequential sampling plans can generally be devised such that each affords the same degree of protection or has nearly identical O.C. curves. For convenience in the discussion that follows, an attributes test is assumed whereby a sampled maintenance task is categorized to be either a success or a failure according to whether the maintenance time is less than or greater than some specified number. For example, such a criterion might be used to test whether 80 percent

**TABLE IX**  
**COMPARISON BETWEEN ATTRIBUTES TEST**  
**AND VARIABLES TEST**

Factor	Attributes Test	Variables Test
Type of Information Yielded	Number or percent of sample that meets some specified characteristic	Observed distribution of some quantitative output
Type of Maintainability Goal	Median or percentile most commonly used	Mean, median, percentile, and variance are most common
Sample-Size Requirements	Higher than variables test for corresponding plan	Lower than for attributes test for corresponding plan
Ease of Application	Data recording and analysis relatively simple	More clerical and analysis costs than for attribute plans
Statistical Considerations	Applies to both parametric and nonparametric tests	Requires an assumption on the underlying distribution unless large sample properties are assumed

of all maintenance actions take less than 20 minutes -- a binomial-type test in which the hypothesis is that 20 minutes is the 80th percentile.

In single sampling, one sample of  $n$  items is tested. Accept or reject decisions are made on the basis of the results by comparing the number of observed unacceptable maintenance actions (i.e., one that takes longer than 20 minutes) with a predetermined acceptance number,  $c$ . In multiple sampling, more than one sample may be necessary before a decision is

reached, but the maximum number of samples and thus the maximum number of items to be tested is known. An example is a double sample plan with the following test criteria:

$$\begin{aligned}n_1 & \text{ (1st sample size) } = 20 \\c_1 & \text{ (accept number for first sample) } = 3\end{aligned}$$

$$\begin{aligned}n_2 & \text{ (2nd sample size) } = 40 \\c_2 & \text{ (accept number for both samples) } = 7\end{aligned}$$

A first sample of 20 items is taken. If 3 or fewer unacceptable maintenance actions are found, an accept decision is made. If 8 or more unacceptable maintenance actions are found, rejection takes place. If 4 to 7 unacceptable maintenance actions are found on the first sample, a second sample of 40 items is taken, and an accept decision is made if the total number of unacceptable maintenance actions is 7 or fewer.

Sequential sampling is an extension of multiple sampling in that decision to accept, reject, or sample further can be made after each individual item (or possibly group of items) is tested. For a standard sequential plan, no maximum number of sample items is specified, although the probability of very large samples is usually quite small. The decision criteria of a sequential sampling plan can be presented graphically. Figure 2 illustrates a sequential test based on a binomial distribution where the number of unacceptable maintenance actions is the decision statistic.

As sampling progresses, the number of unacceptable maintenance actions is plotted against the number of items tested. Testing is continued until the plotted step function crosses one of the two decision lines. Since the step function may remain in the continuous testing region for a long period, especially for borderline lots, truncation or stopping rules can be specified so that the effect on the  $\alpha$  and  $\beta$  errors is negligible.

Multiple sampling generally requires less testing than single sampling, and sequential sampling requires less testing than multiple sampling -- because lots with very good or very poor quality will exhibit such characteristics early in the testing and decisions can be made before multiple samples or further samples in a sequential test are required. Since the first sample of a multiple sampling plan is always smaller than a single sample and since decisions on sequential tests can be made after the results are obtained for each test item, such savings in sample size can be extensive. It is emphasized that the exact sample size of multiple or sequential sampling plans is not predetermined but is a function of the quality of the submitted product. The average sample for various levels of incoming quality can be computed, and the results can be plotted to yield an average sample number (ASN) curve.

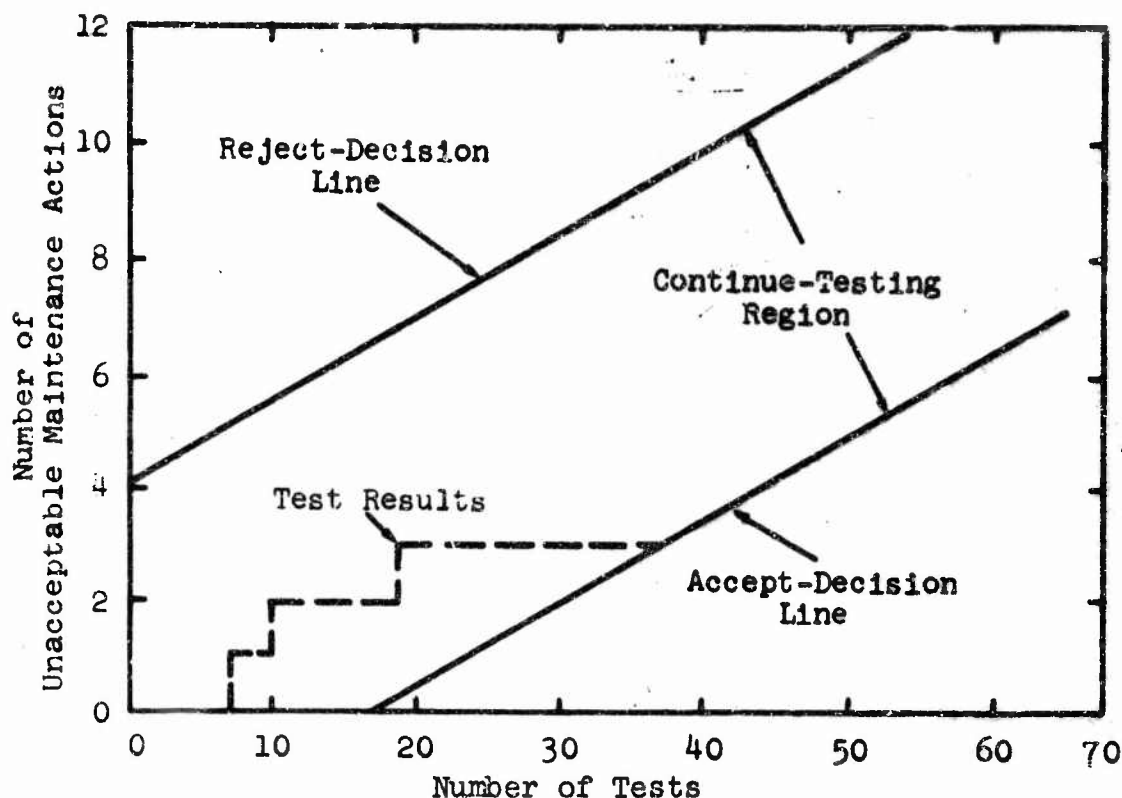


FIGURE 2

#### GRAPHICAL REPRESENTATION OF SEQUENTIAL ACCEPTANCE TEST

An example of these curves is shown in Figure 3 for a maintainability-demonstration test in which the null hypothesis is that 20 minutes is the 95th percentile versus the alternative hypothesis that 20 minutes is the 80th percentile. The  $\alpha$  risk is 0.10, and the  $\beta$  risk is 0.05. The single sampling plan is  $N = 50$ ,  $C = 4$ . Curves for equivalent double and sequential plans are shown in the figure.

Table X is a summary comparison of some characteristics of single, multiple, and sequential sampling plans.

#### 4.6.4 Parametric and Nonparametric Tests

A parametric test is one in which the underlying probability law of the random variable is assumed to take a specific form. In parametric maintainability tests, for example, it is often assumed that maintenance time is a random variable that

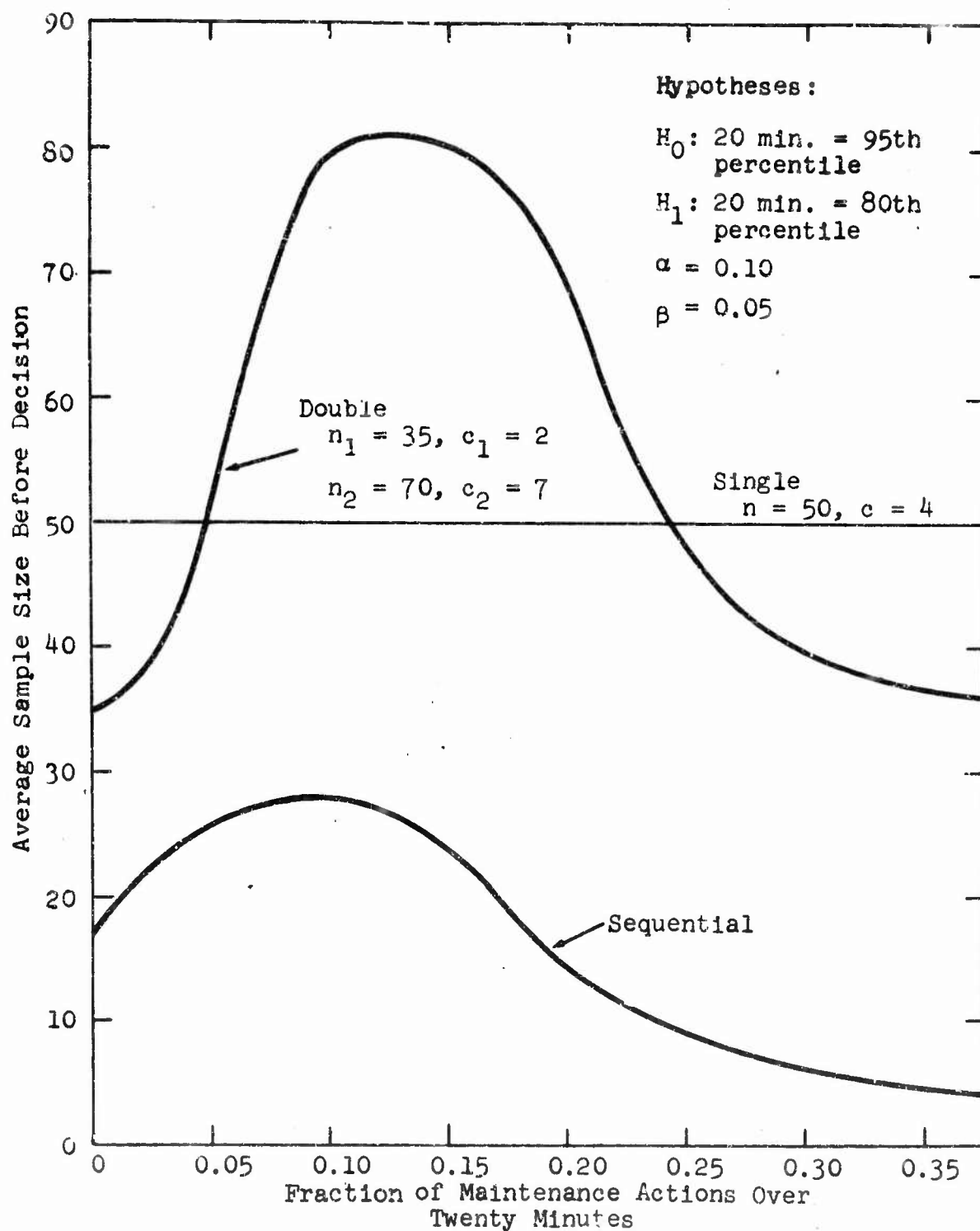


FIGURE 3

AVERAGE AMOUNT OF INSPECTION UNDER SINGLE, DOUBLE, SEQUENTIAL SAMPLING -- BINOMIAL TEST FOR PERCENTILE

TABLE X  
COMPARISON OF SINGLE, MULTIPLE, AND SEQUENTIAL SAMPLING PLANS

Characteristic	Single Sampling	Multiple Sampling	Sequential Sampling
Sample Size	Known	Average can be computed for various incoming-quality levels. Generally less than single.	Average can be computed for various incoming-quality levels. Generally less than single and multiple.
Decision Choices	Accept or reject.	Accept, reject, or take another sample until final sample is selected.	Accept, reject, or test another item.
Predetermined Characteristics	Two of the three quantities $n$ , $\alpha$ , and $\beta$ .	Same as single	Fix $\alpha$ and $\beta$ ; $n$ is a random variable.
Personnel Training	Requires least training.	More trained people required than for single	Requires most training.
Ease of Administration	Easiest. Scheduling can be fairly precise, test-cost estimates can be made.	More difficult than single since the exact number of tests is unknown.	Most difficult in terms of testing, scheduling, and overall administration. Most time-consuming.
Miscellaneous	Best used for testing situations where ease of administration is most important and cost of testing is relatively unimportant.	Has psychological advantage in that supplier is given a "second chance" by taking further samples if first-sample results indicate a marginal lot.	Most efficient test in terms of required sample size. May require approximately 50 percent of sample size of single sampling plans. Best to use when test costs are most important.

can be adequately described by a log-normal distribution. Non-parametric or distribution-free tests are those in which no assumptions about the underlying probability law are made. It is noted that nonparametric tests do not deal with magnitude but with attribute characteristics such as rank, frequency, and ordinal position.

Although the distribution of the attribute tested must be known, it can be inferred without knowing the population distribution of the basic random variable. For example, the number of maintenance-task times (the attribute) less than a constant is a binomial-distributed random variable under some general requirements, irrespective of the distribution of task times (the basic random variable).

Generally, parametric tests are more efficient than non-parametric tests since, for a given amount of testing, more precise estimates or smaller probabilities of incorrect decisions will result than for nonparametric tests. The limitations on the types of parameters that can be tested constitute a disadvantage of nonparametric tests. For example, nonparametric tests of central tendency apply only to the median, while the specifications may be in terms of the mean.

It is emphasized that an incorrect assumption of the underlying probability distribution in a parametric test can lead to an O.C. curve that differs greatly from that planned, especially for small sample sizes. Also, nonparametric tests are generally easy to conduct and evaluate, often requiring only counting, adding, subtracting, or ranking. Because of these two features, nonparametric tests are now receiving much more consideration than in the past.

#### 4.6.5 Classical and Bayesian Tests

A Bayesian test can be generally defined as one that employs prior information in the decision criterion concerning the random variable of interest. The discussion up to this point has been primarily concerned with classical-type tests. It is noted that prior information is used in developing a sampling procedure for stratified sampling, but such use involves test design rather than the decision criterion.

The classical test, as shown earlier, involves a decision criterion based on prescribed probabilities of acceptance for specified maintainability levels. In a Bayesian test, the test results are combined with the prior information to yield a revised (Bayesian) estimate of the actual distribution or parameter, and decisions are made according to the desirability of this estimate.

In essence, if an  $\bar{M}_{ct}$  of one hour is highly desirable, the classical test is designed such that  $P[\text{accept} | \bar{M}_{ct} \leq 1 \text{ hr}]$  is high. A Bayesian test would be one for which  $P[\bar{M}_{ct} \leq 1 \text{ hr} | \text{accept}]$  is high.

Classical tests as defined here are well documented, have been shown to provide the necessary protection against rejecting good or accepting bad product, and thus have been accepted as being a reasonable approach towards assuring product quality.

Bayesian tests are relatively new and their application will therefore have to undergo a trial-and-error and learning process. They do possess two distinct advantages over classical tests: (1) they provide for using available information and therefore have the inherent capability of reducing the test time required before a decision is made; and (2) they can provide assurance on the distribution of outgoing or accepted product, while classical tests generally provide no such control per se.

The major objection to Bayesian tests has been the strong dependence on a prior distribution -- the existence of which some deny and others claim cannot easily be obtained so as to be useful.

The great interest in Bayesian statistics, as evidenced by recent research in statistical theory and applications to maintainability, reliability, and associated disciplines, points to greater use of Bayesian approaches in the future.

## SECTION V

### MAINTAINABILITY-DEMONSTRATION-TEST SPECIFICATION

#### 5.1 GENERAL

A maintainability-demonstration-test specification is defined here as a set of numerical requirements and associated risk levels that will govern the design and decision criteria of the test. For the most common tests, this specification involves decisions regarding the following:

- Type of maintainability index to be specified
- Acceptable and unacceptable values of the index
- Associated risk levels

For example, the test specification might be as follows:

$H_0$ : Mean corrective-maintenance man-hours = 40 minutes

$H_1$ : Mean corrective-maintenance man-hours = 80 minutes

$\alpha = 0.20, \beta = 0.10$

A test based on this specification must be designed such that

$P(\text{reject} \mid \overline{MH}_{ct} = 40 \text{ min}) = 0.20$

$P(\text{accept} \mid \overline{MH}_{ct} = 80 \text{ min}) = 0.10$

The following are some of the more important requisites for a maintainability-demonstration-test specification:

- The maintainability index should represent a measure that is directly influenced by equipment design so that the producer can plan for high assurance of a pass decision, but bears the responsibility for a reject decision.
- Relationships (at least qualitative) between design parameters and the maintainability index should be known so that design evaluations and predictions are possible.
- The maintainability index should be appropriate for, and measurable in, the demonstration-test environment.
- The maintainability index should be related to higher-level system-requirement parameters, and numerical values should be consistent with values for these higher-level parameters.

- Adequate sampling and statistical-evaluation procedures should be available for demonstrating conformance to the requirement.
- Specified maintainability index and risk values should not lead to sample sizes that would exceed available test resources.

Not all of these requisites are necessarily consistent, and often they cannot all be adequately satisfied. A requirement consistent with higher-level goals may result in specified values that require sample sizes larger than expected. Tests for conformance to certain types of requirements may require complex statistical tests that may not be desirable.

It is, therefore, important that the demonstration-test specification be prepared as early as possible so that its implications can be fully evaluated. This will then allow time for a trade-off analysis between test costs and risks of incorrect decisions.

Further details on the three major factors of a maintainability-demonstration specification are discussed in the remainder of this section.

## 5.2 TYPE OF SPECIFIED MAINTAINABILITY INDEX

There are many different types of indices that can be specified for a maintainability demonstration. Some of the more usual alternatives for three major factors are as follows:

<u>Factor</u>	<u>Alternatives</u>
Type of Maintenance Action	Corrective maintenance, preventive maintenance, total maintenance
Type of Statistical Measure	Mean, median, variance, percentile
Type of Time Measurement	Equipment downtime, man- hours, man-hours per operating hour

The above listing represents a possible 36 alternatives; two such are mean corrective-maintenance man-hours, and the 95th percentile of equipment downtime due to all types of maintenance. In addition, there may be multiple parameters such as a mean and percentile and specification of higher-level indices that include maintainability such as availability or effectiveness.

The choice of an appropriate form of the maintainability index can, therefore, be a difficult one. It is the purpose of this section to provide some guidelines for selecting an index appropriate for the maintainability-demonstration task.

It is emphasized that one of the basic purposes of a maintainability demonstration is to provide assurance of accepting equipment with satisfactory maintainability characteristics. This assurance cannot be guaranteed unless the definition of "satisfactory maintainability characteristics" is established. This is not usually an easy task. For example, the operational commander would prefer to have his critically needed equipment operationally ready at all times. The base maintenance commander would prefer equipment that does not tax his manpower organization, and maintenance man-hours may be of more importance to him. Cost control may, perhaps, best be achieved with an index of maintenance man-hours per operating hour.

One approach that at first might appear reasonable is to specify several types of indices, such as mean corrective-maintenance time, mean number of corrective-maintenance man-hours, and man-hours per flight hour. However, these indices are related, and it is quite difficult to develop and apply a valid test for all three indices. Even so, the fact that they are related is helpful since the relationships can be used to specify a value for one type of index with fair confidence that an accepted equipment based on a test of this index will be satisfactory with respect to the related index. Some of the more important relationships are reviewed in the following subsections for each of the three factors cited above.

#### 5.2.1 Type of Maintenance Action

A corrective maintenance action is one performed to restore an item to satisfactory condition. A preventive maintenance action is one performed to detect incipient failures or prevent future failures. These are quite general definitions. For example, in practice, for many supposedly corrective-maintenance actions, no trouble is found and, therefore, no true corrective action is performed.

For specification purposes, the distinguishing feature between the two types of maintenance is that, generally, corrective maintenance is unscheduled, while the preventive-maintenance schedule can often be controlled.

Therefore, with respect to an operational requirement such as availability, corrective maintenance is more critical if preventive maintenance can be scheduled so as not to conflict with operational demands. This may not always be true, however. If the Air Force requires overhaul of a jet engine after 600 operating hours, and if no spares are available and the aircraft is

subject to random demand, the preventive maintenance action of overhauling the engine is as critical as a corrective maintenance action. If a spare engine is installed during the overhaul, the importance of preventive-maintenance time with respect to operational needs is diminished.

The frequency and duration of preventive maintenance actions directly affects maintenance-manpower control. Again, since such actions can often be scheduled, the specification of a corrective-maintenance parameter may be more important than specifying preventive maintenance.

The choice of whether separate indices or combined indices (total downtime or man-hours) of maintainability should be used depends on several factors. If corrective maintenance is more important than preventive maintenance, separate indices and separate tests may be preferred. If downtime due to any cause is critical, a total-downtime index may be used.

From the statistical viewpoint, separate tests are preferred since the distributions of the two types of actions might be different and combining both types would result in a mixture of two distributions, which hinders development of an appropriate test.

There is generally a positive correlation between the statistics for preventive maintenance and those for corrective-maintenance. Many of the tasks are identical (e.g., the final-checkout routine), and the factors that represent good or poor maintenance characteristics will generally affect both types of maintenance action in the same manner. For example, if poor accessibility is a major contributor to excessive corrective-maintenance time on an equipment, it will also adversely influence the preventive maintenance action.

### 5.2.2 Statistical Measures

The mean and median are central-tendency parameters, the variance is a measure of spread, and the percentile specification provides a control on extremes. Mathematically, if  $x$  represents the random variable of interest and  $f(x)$  its continuous probability-density function, the following are defining relationships:

$$\text{Mean: } E(x) = \int_{-\infty}^{\infty} xf(x)dx$$

$$\begin{aligned} \text{Median: } \tilde{M}(x) &= \text{the value of } x \text{ for which } \int_{-\infty}^{\tilde{M}(x)} f(x)dx = \frac{1}{2} \\ &= \int_{\tilde{M}(x)}^{\infty} f(x)dx \end{aligned}$$

$$\text{Variance: } V(x) = E [x - E(x)]^2 = \int_{-\infty}^{\infty} [x - E(x)]^2 f(x) dx$$

$(1-p)^{\text{th}}$  Percentile:  $X_p$  = the value of  $x$  for which

$$\int_{X_p}^{\infty} f(x) dx = p$$

For symmetrical distributions, the mean and median coincide. The median is also the 50th percentile. An indirect control on the variance can be provided by a two-parameter specification such as the mean and 95th percentile or the median and 90th percentile. To demonstrate this, the normal distribution with the following density is considered:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2} (x-\mu)^2}$$

where  $\mu$  is the mean and  $\sigma^2$  is the variance. The  $(1-p)^{\text{th}}$  percentile of the normal distribution is given by

$$X_p = \mu + Z_p \sigma$$

where  $Z_p$  is a standardized normal deviate corresponding to the  $(1-p)^{\text{th}}$  percentile of a normal distribution with mean 0 and variance 1. For  $p = 0.05$ ,  $Z_p = 1.645$ . Thus if the mean and the 95th percentile were specified to be 1 and 3, respectively, the following would be obtained:

$$\mu = 1.0$$

$$X_{0.05} = 3.0$$

Then, from the definition of  $X_{0.05}$ ,

$$X_{0.05} = \mu + Z_p \sigma$$

$$\text{or } 3.0 = 1.0 + 1.645\sigma$$

$$\text{or } \sigma = \frac{2.0}{1.645} = 1.22$$

The distribution most applicable to maintenance times has been found to be the lognormal distribution.<sup>1</sup> The density function is

$$f(x) = \frac{1}{\sqrt{2\pi} \sigma x} e^{-\frac{1}{2\sigma^2} (\ln x - \theta)^2}, \quad x > 0$$

If a random variable has a lognormal distribution function, its logarithm is normally distributed with mean  $\theta$  and variance  $\sigma^2$ . The following relationships pertain to the lognormal:

$$\text{Mean: } E(x) = e^{\theta + \sigma^2/2}$$

$$\text{Median: } \tilde{M}(x) = e^{\theta}$$

$$\text{Variance: } V(x) = e^{2\theta + \sigma^2}(e^{\sigma^2} - 1)$$

$$(1-p)^{\text{th}} \text{ percentile: } x_p = e^{\theta + Z_p \sigma} \quad [Z_p = (1-p)^{\text{th}} \text{ percentile of normal (0,1) distribution.}]$$

Because of these relationships, the specification of any two of the above parameters completely defines the distribution (see Table XXXII, Section VII). The fact that the median is independent of  $\sigma^2$  is one reason why this parameter is often associated with the lognormal since it allows the application of relatively simple tests.

A detailed study of the relationships of lognormal parameters and the specification of such parameters is presented in the Rome Air Development Center Technical Report, RADC-TR-67-403, Maintainability Parameters and Their Relationships, J. Klion, September 1967.

### 5.2.3 Type of Time Measurement

Equipment downtime is the time measurement most related to operational requirements. Man-hours and man-hours per operating hour are more closely associated with manpower and cost control although they are, of course, related to downtime.

Average man-hours per maintenance action can be estimated by multiplying the average number of maintenance men per maintenance action by the average downtime.

<sup>1</sup>See Appendix A for a detailed discussion of the lognormal distribution.

For 13 equipments monitored in this study, estimates of total active-maintenance man-hours (excluding "no trouble found" actions) based on this relationship yielded good results. The estimated and observed man-hours, along with the absolute value of the relative error, are shown in Table XI. The average of the relative errors is 7.6 percent.

TABLE XI  
OBSERVED AND ESTIMATED MAINTENANCE TIME  
(Maintenance Man-Minutes)

Equipment	Observed	Estimated	Absolute Relative Error*
A04	60.3	63.0	0.046
A08	282.6	322.8	0.142
A10	174.6	162.4	0.070
B04	118.8	106.8	0.101
B06	61.5	57.8	0.061
B09	81.1	78.5	0.032
B10	98.1	94.6	0.036
B12	71.2	79.5	0.118
B13	20.3	22.5	0.110
B14	71.0	74.0	0.043
D01	155.8	153.2	0.017
D02	77.9	65.3	0.162
D03	101.5	114.6	0.129
* $\frac{ \text{Observed} - \text{Estimated} }{\text{Observed}}$			

For complete systems such as aircraft, for which concurrent maintenance actions can and frequently do take place, the relationship between equipment downtime and system downtime provides a basis for allocating an overall system-downtime requirement. If a system is divided into n equipments such that (1) concurrent maintenance can take place only on different equipments and not within an equipment, and (2) the probability that more than

two equipments will require concurrent maintenance is negligible, then the following relationship can be used as a model for average system downtime based on equipment downtimes:

$$\bar{t}_s = \sum_{j=1}^n \sum_{i=1}^{n_j} p_{ij} \bar{t}_{ij} + \sum_{j=1}^{n-1} \sum_{i=1}^{n_j} \sum_{\ell=j+1}^n \sum_{k=1}^{n_\ell} p_{ij,k\ell} \cdot \max [\bar{t}_{ij}, \bar{t}_{k\ell}]$$

where

$\bar{t}_s$  = average system downtime

$n_j$  is the number of possible tasks for the  $j^{\text{th}}$  equipment

$p_{ij}$  is the probability that only the  $i^{\text{th}}$  task in the  $j^{\text{th}}$  equipment is performed given system failure

$p_{ij,k\ell}$  is the probability that the  $i^{\text{th}}$  task in the  $j^{\text{th}}$  equipment and the  $k^{\text{th}}$  task in the  $\ell^{\text{th}}$  equipment are required concurrently given system failure

$\bar{t}_{ij}$  is the average time for completing the  $i^{\text{th}}$  task in the  $j^{\text{th}}$  equipment

Three models for estimating system man-hours from downtime estimates are shown below. The following notation is used:

$MH_s$  = System man-hours

$\bar{m}_s$  = Average number of men per system maintenance action

$\bar{m}_s(r)$  = Average number of men per system maintenance action when  $r$  failures are involved

$\bar{m}_{ij}$  = Average number of men involved when the  $i^{\text{th}}$  maintenance task is required for the  $j^{\text{th}}$  equipment

$\bar{t}_s$  = Average system downtime

$\bar{t}_s(r)$  = Average system downtime when  $r$  failures occur

$\bar{t}_{ij}$  = Average downtime of the  $j^{\text{th}}$  equipment when the  $i^{\text{th}}$  maintenance task is to be performed

$p_r$  = Probability that  $r$  maintenance tasks are required given system failure

- $p_{1j}$  = Probability that only the  $i^{\text{th}}$  maintenance task in the  $j^{\text{th}}$  equipment is required for a system failure
- $p_{1j,kl}$  = Probability that the  $i^{\text{th}}$  maintenance task for the  $j^{\text{th}}$  equipment and  $k^{\text{th}}$  maintenance task for the  $l^{\text{th}}$  equipment are required for a system failure

The models are as follows:

- A. Model based on average system downtime:

$$E(MH_S) = \bar{m}_S \bar{t}_S$$

- B. Model based on average system downtime as a function of number of equipment failures:

$$E(MH_S) = \sum_r p_r \bar{m}_S(r) \bar{t}_S(r)$$

- C. Model based on average equipment downtime assuring no more than two concurrent tasks:

$$E(MH_S) = \sum_j \sum_i p_{ij} \bar{m}_{1j} \bar{t}_{1j} + \sum_j \sum_i \sum_l \sum_k p_{1j,kl} [\bar{m}_{1j} \bar{t}_{1j} + \bar{m}_{kl} \bar{t}_{kl}]$$

Models A and B are both based on system-downtime statistics. The former requires only an overall estimate of system downtime and manning, while the latter requires estimates of downtime and manning as a function of the number of concurrent tasks required. Since manpower is generally limited, if a large number of simultaneous failures requiring many concurrent maintenance tasks occur, system downtime may be partially due to the unavailability of maintenance men. This factor is accounted for directly in Model B through the  $\bar{t}_S(r)$  variable but only indirectly in Model A.

Model C uses equipment rather than system-downtime values directly. While these may be more readily available, no accounting is made for total system checkout after all tasks are completed, although an appropriate "K" factor can easily be incorporated. While the equation shown limits the number of concurrent tasks to two, it is apparent that an extension can be made for more than two tasks.

A model for the expected value of man-hours per operating hour (MH/OH), showing its relationship to average downtime, can be developed as follows:

$$\begin{aligned}
 E[\text{MH/OH}] &= \frac{1}{T} \sum_{j=0}^{\infty} P[j \text{ failures in } T \text{ operating hours}] E[\text{MH} | j \text{ failures}] \\
 &= \frac{1}{T} \sum_{j=0}^{\infty} P[j \text{ failures in } T \text{ operating hours}] j E[\text{MH} | 1 \text{ failure}] \\
 &= \frac{E[\text{MH} | 1 \text{ failure}]}{T} \sum_{j=0}^{\infty} j P[j \text{ failures in } T \text{ operating hours}] \\
 &= \frac{E[\text{MH} | 1 \text{ failure}]}{T} E[\text{number of failures in } T \text{ operating hours}]
 \end{aligned}$$

If a constant equipment failure rate is assumed, the following rate is obtained:

$$E[\text{MH/OH}] = \frac{E[\text{MH/action}]}{T} \lambda T = \lambda E[\text{MH/action}]$$

Again, by using the relationship

$$E[\text{MH/action}] = E[\text{number of men/action}][\text{average downtime}] = \bar{m} \bar{t}$$

the following is obtained:

$$E[\text{MH/OH}] = \lambda \bar{m} \bar{t}$$

This general expression applies at either the equipment level or the system level, but it is more accurate for the former because the simplifying equations do not account directly for concurrent maintenance. It is emphasized that the man-hour rate is directly influenced by reliability. From the equation

$$E[\text{MH/OH}] = \lambda \bar{m} \bar{t}$$

it is seen that the man-hour rate changes in direct proportion to changes in the failure rate,  $\lambda$ . This is not true for average downtime or average man-hours per maintenance action, since the reliability factor for these measures influences only the relative frequency of the various maintenance tasks given that a maintenance action is required.

Since a demonstration test based on man-hour rate is a test that includes both reliability and maintainability factors, it does not truly fall in the category of a maintainability-demonstration test in which the maintainability-design group bears the major responsibility for success in meeting goals. This is not to imply that a test based on man-hour rate is not useful; however, it is important to recognize the influences on this type of measure so that appropriate responsibilities and control can be established.

Because of the direct relationship between expected man-hours per maintenance action and man-hour rate, and the fact that control of the latter is as much a responsibility of reliability design as of maintainability design, man-hour rate as a maintainability index was not treated in as much depth in this study as the downtime and man-hour indices.

The relationships shown above that relate man-hours and man-hour rates to equipment downtimes do show that the specification of equipment downtimes does permit evaluation of man-hour parameters at either the system or equipment level. The complexity of the relationships for system-level requirements, however, does highlight the problems that can occur if a system man-hour requirement is to be tested by synthesizing results obtained at the equipment level.

#### 5.2.4 Guidelines for Index Selection

The principal objective in selecting the index for a maintainability-demonstration test should be to seek the one that is most consistent with the mission objectives and operational constraints. Generally, this will mean that equipment downtime is the time measurement of the index since operational effectiveness cannot be achieved unless downtime is controlled.

If the need for an equipment is not critical, and manpower control is important, a man-hour index may be most appropriate. Preventive-maintenance man-hours per operating hour is preferable to downtime due to preventive maintenance for equipments for which such maintenance can be scheduled without fear of operational demand during the maintenance action.

By the same reasoning, corrective maintenance is more crucial than preventive maintenance, especially if the latter can be scheduled to take place during known periods of non-use. For continuously needed equipment, such as an alert radar, total maintenance time is of prime importance. For equipment demanded at random times, such as a missile-defense equipment, the approach might be to use separate controls for corrective maintenance and preventive maintenance. The choice of the statistical measure to be used often depends on the mission objective. If there is an

availability requirement for the system, the following relationship is often used to determine reliability and maintainability requirements:

$$\text{Availability} = \frac{\text{MTBF}}{\text{MTBF} + \text{MDT}}$$

This relationship provides a basis for trade-off between MTBF and MDT. Several such trade-off curves are shown in Subsection 5.3.2 for various availability requirements.

When this availability expression is appropriate, a mean value should be used for the maintainability index. However, there may be an availability requirement, for which a maximum downtime is more appropriate. Such a requirement would apply to critical equipment aboard an aircraft where the aircraft may have to be available for a new mission within two hours after completing a mission. In this case a requirement of, say, 0.95 probability of completing any necessary maintenance within 100 minutes would be more consistent with the operational objective than a mean-value index. This is discussed in greater detail in Subsection 5.3.

Another element to be considered in the choice of the statistical measure is the underlying distribution of maintenance times. For symmetrical distributions, the mean and median are identical, and the choice depends solely on statistical niceties. For the more common skewed distributions of maintenance time, the mean is strongly influenced by the long maintenance times but the median is not. When either can be used, the mean generally provides better manpower cost control, is derivable from a higher-level specification, and has more desirable statistical properties (e.g., application of the central-limit theorem). The median is applicable to distribution-free tests, has direct operational meaning in the sense that 50 percent of all maintenance actions will be performed within the median-time period, and for the log-normal distribution is dependent on only one parameter as contrasted with two for the mean.

The foregoing discussion has provided some direction on choosing an appropriate maintainability index for several situations. To provide greater detail and permit more definitive recommendations, an established matrix for selecting maintainability measures is used here.<sup>2</sup> In this procedure, seven conditions that should influence the choice of a measure or index are listed, and recommended indices for common combinations of such conditions are given in the matrix. The conditions and matrix presented here are modified slightly from those of the Notebook to make them applicable to demonstration.

<sup>2</sup>This matrix will appear in the "Maintainability Engineering Design Notebook", being prepared by ARINC Research Corporation for Rome Air Development Center under Contract F30602-68-C-0208.

The matrix is presented in Exhibit 2. To use the matrix, each of the conditions listed at the top of the exhibit that apply to the equipment of interest should be checked. The appropriate index is then found from the matrix by locating the column that contains an x for each condition checked above. For example, if steady-state availability is a critical parameter (Condition 1) and maintenance time is limited by environmental or operational circumstances (Condition 5), the recommended index provides a control on both the mean and maximum maintenance time, and there is an option for including preventive-maintenance time depending on equipment use or scheduling and criticality.

The set of conditions listed is not exhaustive, but it is believed to include the most important ones.

Several of the major considerations that led to the development of the matrix are as follows:

- The mean is directly related to steady-state availability and is therefore the index of choice when this operational requirement exists.
- If the distribution of maintenance times is unknown, the median is preferred since it permits distribution-free tests. If availability is critical, however, use of the central-limit theorem permits a mean test provided the sample size is large.
- For the lognormal distribution, the median is preferred to the mean (assuming that Condition 2 applies and that 5 and 6 do not) since it is based on only one parameter, which makes statistical analysis exact.
- When maintenance time is limited (Condition 5), the  $M_{\max}$  index is preferred.
- The mean is preferred over the median if manpower control is also required because the mean is more directly related to man-hours. However, if the distribution is unknown, the median may be used as long as availability is not critical.

Complete dependence on this procedure is to be avoided. Because of the wide variety of equipments, mission objectives, and environmental and operational circumstances, the selection matrix should be considered a guide only. Ultimately, the best measure is determined by individual system circumstances and good judgment.

**Condition Identification**  
(Place X in appropriate boxes)

**Condition**

- ☐ 1 Steady-state availability is a critical parameter.
- ☐ 2 Steady-state availability is not a critical parameter.
- ☐ 3 Maintenance-time distribution is unknown.
- ☐ 4 Maintenance-time distribution is expected to be lognormal.
- ☐ 5 Environmental or operational circumstances limit maintenance time.
- ☐ 6 Manpower allocation or cost is an important factor.

Selection Matrix <sup>①</sup>																			
<div><div><math>\bar{M}</math> Index</div><div>Condition</div></div>	$\bar{M}_{ct}$ and $\bar{M}_{pt}$ <sup>②</sup>				$\tilde{M}_{ct}$			$M_{max}$			MMH			$\bar{M}_{ct}$ and $M_{max\ ct}$ $\bar{M}_{pt}$ and $M_{max\ pt}$ <sup>②</sup>				$\tilde{M}_{ct}$ and $M_{max\ ct}$ $\tilde{M}_{pt}$ and $M_{max\ pt}$	
1	X	X	X	X											X	X	X	X	
2					X	X	X	X	X	X	X	X	X					X	X
3			X			X				X			X			X			X
4			X	X			X							X			X	X	
5								X	X	X					X	X	X	X	X
6				X								X	X	X				X	X

**Notation**  $\bar{M}$  = mean,  $\tilde{M}$  = median,  $M_{max}$  = maximum maintenance time, MMH = maintenance man-hours (percentile)

ct = corrective maintenance, pt = preventive maintenance

**Notes**

- ① The inclusion of preventive-maintenance indices is optional depending on scheduling and criticality.
- ② A combined total-maintenance-time index can be used instead of separate indices for corrective and preventive maintenance.

**EXHIBIT 2**

**PROCEDURE FOR MAINTAINABILITY-INDEX SELECTION**

### 5.3 SPECIFIED VALUES FOR THE MAINTAINABILITY INDEX

#### 5.3.1 Three Basic Criteria

As discussed in Section 4.3, the usual specification of values for maintainability demonstration involves assignment of two values for the index selected -- a desirable value associated with the null hypothesis,  $H_0$ , and an undesirable (sometimes called marginally acceptable) value associated with the alternative hypothesis,  $H_1$ .

In assigning such values, it is reasonable first to consider the goal or  $H_0$  value, since this is what the producer and consumer both seek, and then to assign the  $H_1$  value, which will be a function of the desirable value, minimum operational goals, and other factors such as required sample sizes.

There are three basic criteria for specifying the desirable values of the selected maintainability index:

- (1) The specified value should be consistent with higher-system-level requirements.
- (2) It should be realistic.
- (3) It should pertain to the demonstration environment.

Unfortunately, the first two criteria are sometimes in conflict because higher-level requirements may be unrealistic. As a minimum, any maintainability-index value derived from higher-level requirements should be checked for reasonableness. If the derived values are not reasonable (exceed the state of the art), the higher-level goal is suspect. If this goal cannot be compromised, however, the producer can be given some reprieve by adjustment of risk values.

On the other hand, if a derived value is less acceptable than what can normally be expected, the values should be readjusted to state-of-the-art levels.

The third criterion may also be a source of conflict. A higher-level requirement such as overall system effectiveness may be translated into capability and availability requirements by such means as a WSEIAC-type analysis. The availability must then be further refined to reliability and maintainability indices. The operational environment to which the maintainability index applies, however, may be significantly different from the demonstration environment under which the equipment or system is to be tested.

It is, naturally, preferable to conduct the test in the actual operational environment, but this is often impossible. If the test environment does differ from the operational environment, it is reasonable to adjust a maintainability goal based on operational need to reflect the differences between test and field conditions.

### 5.3.2 Specification Based on Higher-System-Level Requirements

There are many types of system-level requirements that are directly or indirectly related to the maintainability characteristics of a system. The most important of these system operational-type requirements is availability, which in the most general sense is a measure of the readiness of the system for operational use on demand. Two of the more important types of availability measures are as follows:

- Point Availability - The probability that the system is available for operational use at a random point in time.
- Interval Availability - The probability that the system will be available for operational use within a specified time interval.

Point availability is generally applicable to systems whose mission is continuous, such as an alert radar. For these types of systems, the ratio of on time to total time must be high, and this ratio is best expressed by the general steady-state availability expression

$$A = \frac{MTBM}{MTBM + MDT}$$

where

MTBM = mean time between maintenance

MDT = mean downtime

When preventive or noncorrective maintenance can be scheduled so that it does not conflict with mission objectives, the following expression is applicable:

$$A = \frac{MTBF}{MTBF + MTTR}$$

where

MTBF = mean time between failures

MTTR = mean time to repair

Interval availability as defined above is applicable when the system is required to perform a series of missions, the most common example of such a system being an aircraft. For such cases, it is often important to control the probability of readiness after completion of a mission by an interval-availability requirement.

A model for this type of requirement can be fairly complex depending on the system, operational conditions, and assumptions made. A relatively simple model for steady-state interval availability -- assuming a Markov process for the mission/service-repair sequence, constant mission time  $T$ , and constant allowable repair time  $t$  -- is presented below.

Let

$A(t)$  represent the probability that the system is available within  $t$  hours after scheduled mission completion

$R(T)$  represent the system reliability for a mission of  $T$  hours

$S(t)$  represent the probability that necessary servicing (e.g., refueling and rearming an aircraft) is performed within  $t$  hours after a successful mission

$M(t)$  represent the probability that servicing and any necessary repairs can be accomplished within  $t$  hours after initiation of maintenance on a failed system

The steady-state interval availability is then given by the following equation [a bar above a symbol represents the complementary event, e.g.,  $\bar{R}(t) = 1 - R(t)$ ]:

$$A(t) = A(t)R(T)S(t) + A(t)\bar{R}(T)M(t) + \bar{A}(t)M(T + t)$$

The first term on the right-hand side is the probability that the system was available at the start of the previous mission, did not fail in  $T$  hours of operation, and is serviced within  $t$  hours. The second term represents the probability that the system was available at the start of the previous mission, a failure occurred during that mission, and repair and servicing are completed within  $t$  hours. The third term is the probability that the system was unavailable at the start of the previous mission and repair and servicing is completed before the start of the current mission (a total time of  $T+t$  hours).

Solving for  $A(t)$  yields

$$A(t) = \frac{M(T+t)}{1 - R(T)S(t) - \bar{R}(T)M(t) + M(T+t)}$$

The assumptions of constant mission and allowable maintenance times, can be relaxed, and such factors as malfunction-detection probability and repair efficiency can be included at the expense of additional model complexity. For illustrative purposes, however, the above model will be retained.

By using the preceding equations that relate availability to maintainability characteristics, it is possible to determine maintainability requirements from an availability requirement. Since reliability factors are also involved, such determination is best made through a trade-off process wherein feasibility and costs are also considered in selecting the appropriate set of reliability and maintainability requirements. The three cases discussed above are considered with respect to such trade-offs in the following subsections.

#### 5.3.2.1 Point Availability Including Preventive Maintenance

The availability equation for point availability including preventive maintenance is as follows:

$$A = \frac{MTBM}{MTBM + MDT}$$

To obtain MTBM (mean time between maintenance), both preventive and corrective maintenance must be considered. It will be assumed that the mean time between failures (MTBF) is equal to  $\theta$  and that preventive maintenance is scheduled every  $T_p$  hours.

With failure time denoted by  $t_f$ ,

$$MTBM = P[t_f < T_p] E[t_f | t_f < T_p] + P[t_f \geq T_p] T_p$$

If it is assumed that an exponential assumption is adequate for describing the failure-time distribution, then

$$\begin{aligned} MTBM &= \frac{\left(1 - e^{-T_p/\theta}\right) \int_0^{T_p} t_f \frac{1}{\theta} e^{-t_f/\theta} dt_f}{1 - e^{-T_p/\theta}} + e^{-T_p/\theta} T_p \\ &= \theta [1 - e^{-T_p/\theta} (1 + T_p/\theta)] + e^{-T_p/\theta} T_p \\ &= \theta \left(1 - e^{-T_p/\theta}\right) \end{aligned}$$

If  $T_p/\theta$  is small, say  $T_p/\theta \leq 0.05$ , then  $1 - e^{-T_p/\theta} \approx T_p/\theta$  and  $MTBM \approx T_p$ , which is reasonable since preventive maintenance actions occur much more frequently than corrective maintenance actions. Conversely, if  $T_p/\theta$  is large, say  $T_p/\theta \geq 3$ , then  $1 - e^{-T_p/\theta} \approx 1$  and  $MTBM \approx \theta$ , which, again, is reasonable since corrective maintenance occurs much more frequently than preventive maintenance.

The mean downtime (MDT) parameter can be estimated as follows: In  $T$  total hours,  $T/T_p$  preventive maintenance (PM) actions and  $T/\theta$  corrective maintenance (CM) actions can be expected to take place. The probability that a maintenance action is preventive is then

$$P [PM] = \frac{T/T_p}{T/T_p + T/\theta} = \frac{\theta}{\theta + T_p}$$

and, similarly,

$$P [CM] = \frac{T_p}{\theta + T_p}$$

Then

$$MDT = \frac{\theta}{\theta + T_p} \bar{M}_{pt} + \frac{T_p}{\theta + T_p} \bar{M}_{ct}$$

Then, for point availability,

$$A = \frac{\theta(1 - e^{-T_p/\theta})}{\theta(1 - e^{-T_p/\theta}) + \left(\frac{1}{\theta + T_p}\right)(\theta \bar{M}_{pt} + T_p \bar{M}_{ct})}$$

Of particular interest for maintainability demonstration is a choice of values for  $T_p$ ,  $\bar{M}_{pt}$ , and  $\bar{M}_{ct}$  given a requirement on  $A$ . If the time interval between preventive maintenance actions ( $T_p$ ) is increased, it might be reasonable to lengthen  $\bar{M}_{pt}$  since the tasks may be more extensive as a result of the longer operating time. Also,  $\theta$  may be adversely affected if  $T_p$  is made too long. On the other hand, too small a value for  $T_p$  increases the number

of downtimes due to preventive maintenance; and while  $\theta$  may be increased somewhat and  $\bar{M}_{pt}$  decreased, there is a minimum  $T_p$  value below which it would be unwise to specify.

A general trade-off relationship is difficult to develop because the interrelationships that exist may be varied and complex. Instead, a simple numerical example is provided here.

Assume that there is an availability requirement of 0.96. From past experience, feasibility analyses, and operational requirements, the following are reasonable ranges for the parameters listed:

$$\begin{array}{ll} \theta = 50 - 150 \text{ hours} & \bar{M}_{pt} = 1 - 3 \text{ hours} \\ T_p = 25 - 75 \text{ hours} & \bar{M}_{ct} = 1 - 4 \text{ hours} \end{array}$$

If the worst extreme is considered, i.e.,  $\theta = 50$ ,  $T_p = 25$ ,  $\bar{M}_{pt} = 3$ ,  $\bar{M}_{ct} = 4$ , then

$$A = \frac{1 - e^{-1/2}}{1 - e^{-1/2} + \frac{1}{75}(3+2)} = \frac{0.61}{0.61 + 0.067} = 0.91,$$

indicating that the goal cannot be met without careful attention to requirements. On the other hand, under the best-case conditions of  $\theta = 150$ ,  $T_p = 75$ ,  $\bar{M}_{pt} = \bar{M}_{ct} = 1$ ,

$$A = \frac{1 - e^{-1/2}}{1 - e^{-1/2} + 1.5/225} = \frac{0.61}{0.61 + 0.007} = 0.99,$$

indicating that the goal is feasible with an appropriate set of requirements.

Assume now that a more detailed analysis between  $T_p$ ,  $\theta$ , and  $\bar{M}_{pt}$  yields the following alternatives:

<u>Alternative</u>	<u><math>T_p</math></u>	<u>Max <math>\theta</math></u>	<u>Min <math>\bar{M}_{pt}</math></u>
I	25	150	1.0
II	50	100	1.5
III	75	75	2.0

The values of  $\bar{M}_{ct}$  that provide an availability of 0.95 are determined from the following equation:

$$\bar{M}_{ct} = \frac{T_p + \theta}{r} \left[ \frac{1 - e^{-r}}{A} + e^{-r} - 1 - \frac{\bar{M}_{pt}}{T_p + \theta} \right]$$

where

$$r = T_p / \theta.$$

The results are as follows:

<u>Alternative</u>	<u><math>\bar{M}_{ct}</math></u>
I	0.72
II	1.92
III	1.36

Because of the initial restriction on  $\bar{M}_{ct}$  of  $1 \leq \bar{M}_{ct} \leq 4$ , Alternative I cannot be chosen. Therefore, the choice is between II and III, and this decision would depend on the costs associated with the specific values of  $T_p$ ,  $\theta$ ,  $\bar{M}_{pt}$ , and  $\bar{M}_{ct}$ .

This particular example involves the selection of a preventive-maintenance schedule as well as mean corrective-maintenance and preventive-maintenance times. Much more sophisticated models for preventive-maintenance scheduling have been developed, and in practice the procedure might be to use one of these models to select  $T_p$  and  $\theta$  and then choose values for  $\bar{M}_{ct}$  and  $\bar{M}_{pt}$  to meet the availability goal.

#### 5.3.2.2 Point Availability Excluding Preventive Maintenance

The following availability expression for point availability excluding preventive maintenance can be used to determine reliability and maintainability requirements:

$$A = \frac{MTBF}{MTBF + MTTR}$$

If an MTBF goal has already been established or is quite restricted by the state of the art, the simple relationship

$$MTTR = MTBF \left( \frac{1}{A} - 1 \right)$$

yields the required MTTR value. For the more common case, in

which there are ranges of possible values for both MTR and MTRF, the curves shown in Figure 4 provide a basis for trade-off. In the figure use the left-hand vertical scale with the bottom horizontal scale or the right-hand vertical scale with the top horizontal scale.

### 5.3.2.3 Interval Availability

From the interval-availability expression

$$A(t) = \frac{M(T+t)}{1 - R(T)S(t) + \bar{R}(T)M(t) + M(T+t)}$$

the maintainability parameters of interest are  $M(t)$ ,  $M(T+t)$ , and  $S(t)$ .  $M(T+t)$  should equal 1 since this represents the probability that maintenance is completed within the usual allowable time ( $t$ ) plus the mission time  $T$ . Then

$$A(t) = \frac{1}{2 - R(T)S(t) + \bar{R}(T)M(t)}$$

Since a maximum of  $t$  hours is available for servicing and corrective maintenance, servicing should be completed in much less time than  $t$  hours to permit corrective maintenance to take place. In this case, a time  $t_s < t$  can be chosen such that requirements are to be placed on  $S(t_s)$  and  $M_c(t_c)$ , where  $t_s$  plus  $t_c$  is less than or equal to  $t$ ,  $S(t_s)$  equals the probability that servicing is completed within  $t_s$  hours, and  $M_c(t_c)$  equals the probability that corrective maintenance is completed within time  $t_c$ . Then  $M(t)$  can be replaced by  $S(t_s) \times M_c(t_c)$  (assuming the independence of the two associated events). The use of this product is conservative since it is assumed that only  $t_c$  hours are available for corrective maintenance even if servicing is completed earlier than  $t_s$  hours. The availability model is then

$$A(t) = \frac{1}{2 - R(T)S(t_s) - \bar{R}(T)S(t_s)M_c(t_c)}$$

Trade-off curves relating  $R$ ,  $S$ , and  $M$  to  $A$  are shown in Figure 5. Again, cost and operational factors will determine which of the appropriate combinations of  $R$ ,  $S(t_s)$ , and  $M_c(t_c)$  to specify for a given availability requirement. In this example the  $S(t_s)$  and  $M_c(t_c)$  requirements are often called  $M_{\max}$ -type requirements, which are actually percentile values of the cumulative distribution function.

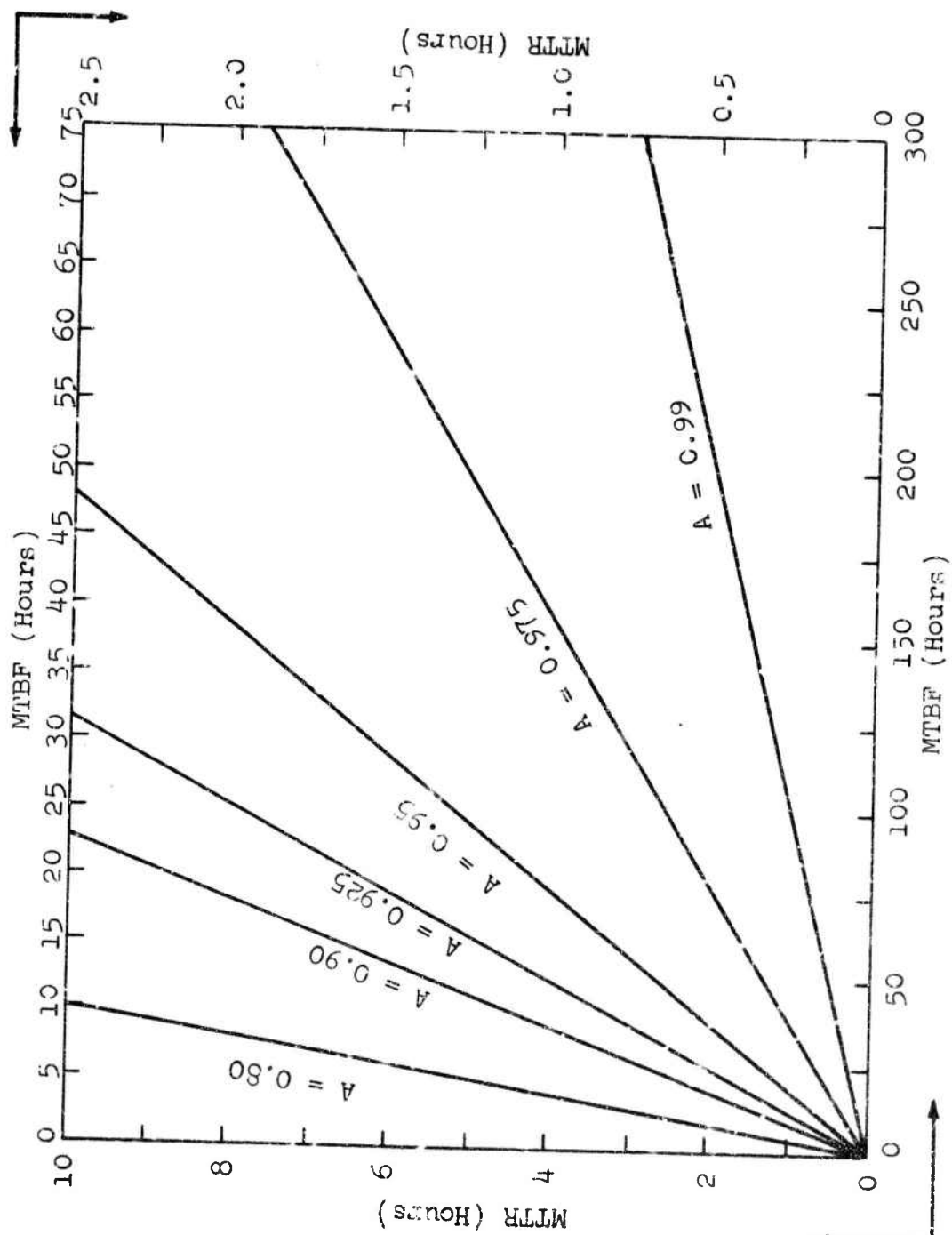


FIGURE 4

RELATIONSHIPS BETWEEN POINT AVAILABILITY,  
MTTR, AND MTBF

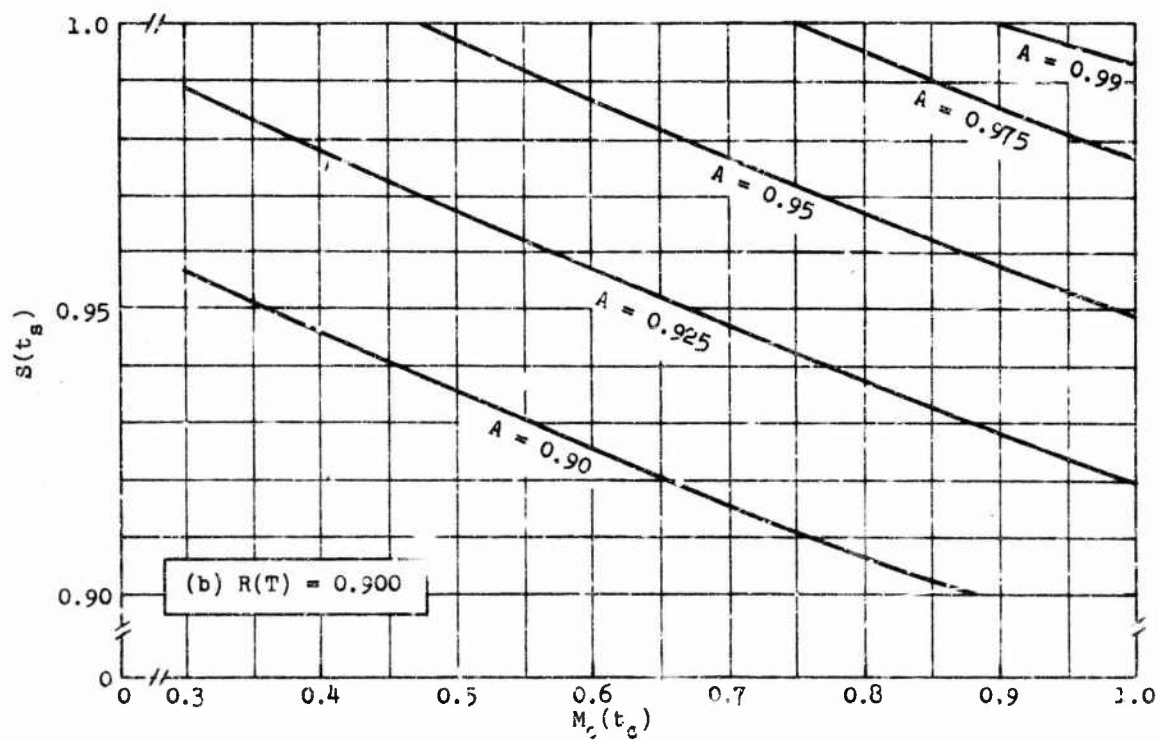
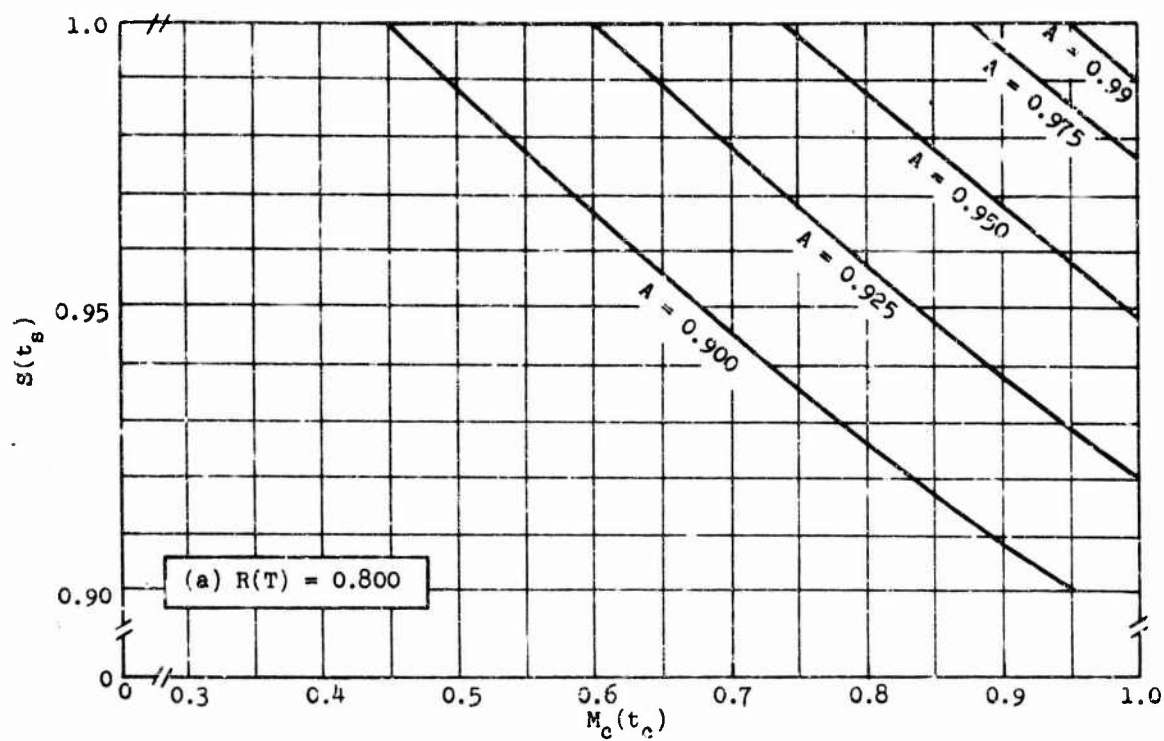


FIGURE 5

TRADE-OFF RELATIONSHIPS BETWEEN  $S(t_s)$  AND  $M_c(t_c)$  TO MEET A SPECIFIED INTERVAL AVAILABILITY REQUIREMENT,  $A$

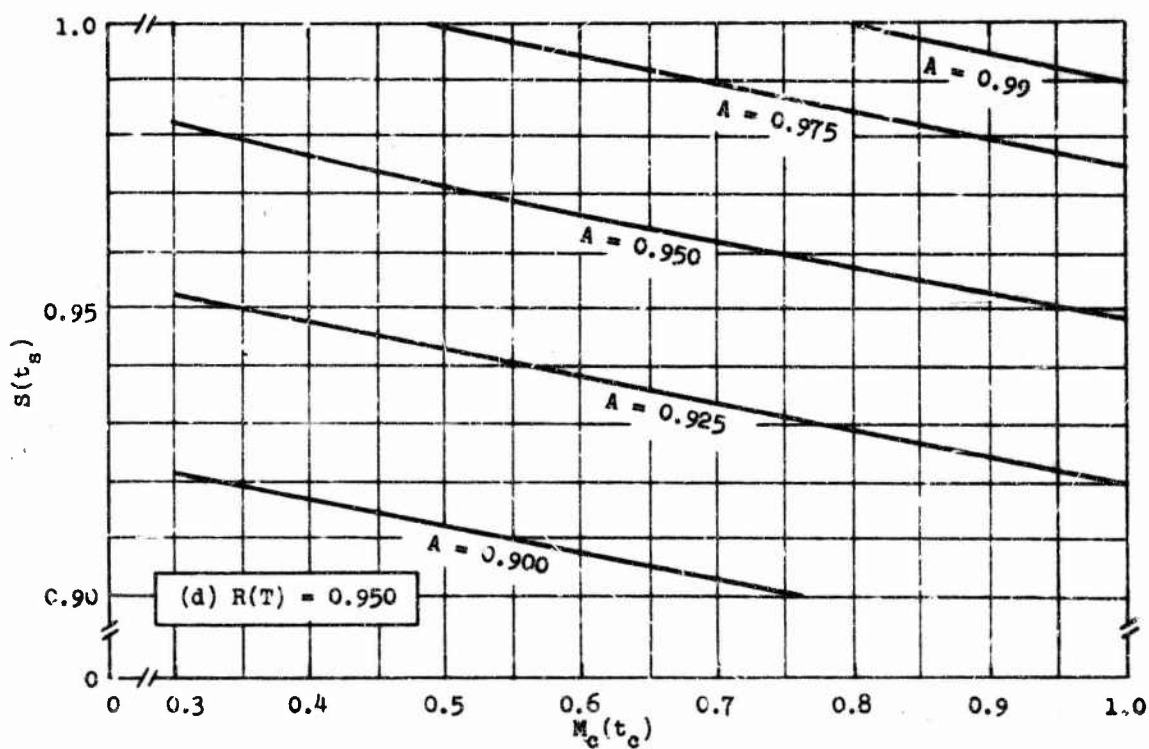
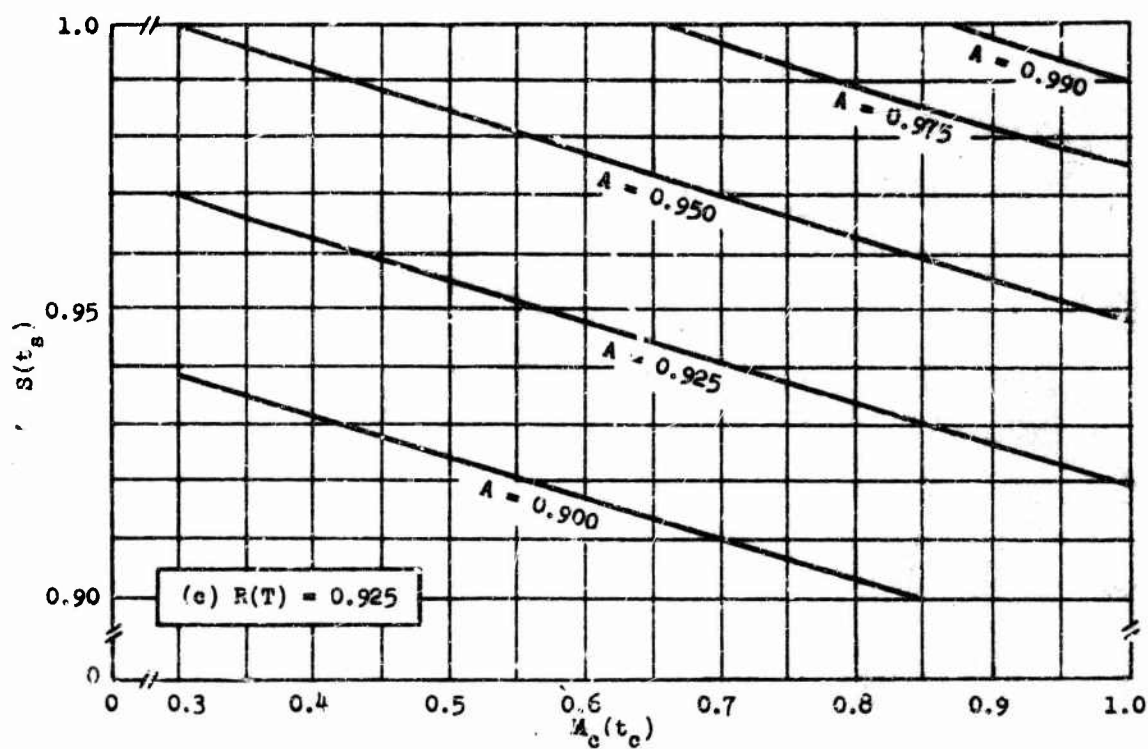


FIGURE 5 --- Concluded

#### 5.3.2.4 Applicability of Approaches

The three above-described approaches for obtaining maintainability requirements from an overall availability requirement are only indicative of the type of approach that can be used. Several simplifying assumptions were made in establishing the relationships, some possibly important factors were not included, and cost was given only qualitative consideration. Therefore, the equations and curves presented for obtaining specified values must be adjusted to account for factors that have not been considered adequately in this general model.

#### 5.3.3 Realism of Specified Values

Approaches similar to those presented in Subsection 5.3.2 lead to a specified maintainability value. The next criterion is one of realism. It is necessary first to establish what is meant by a realistic value. Expressions such as "within the state of the art" are commonly encountered, and while they do not provide a quantitative assessment, they do convey the general belief that that value can be achieved by current technological capability.

Since maintainability-demonstration-test requirements must be established very early in the development program (often before contract award) the most logical approach to assessing realism, and sometimes even establishing the requirement if allocation from higher levels is not required, is to evaluate the maintainability performance of existing systems similar to that under development. If the basic maintainability design is known at the time the requirement is to be established, an applicable prediction technique can be exercised.

Whether historical data or prediction, or both, is used for assessing realism, careful judgment is required. If an allocation leads to an  $M_{ct}$  value of 20 minutes but a 30-minute value was observed for the most similar existing system, can it be concluded that 20 minutes is unrealistic? The following questions must be considered:

- (1) How similar are the items?
- (2) How similar will the maintenance environment be?
- (3) Since the observed 30-minute value is necessarily based on a sample, what is the lower confidence limit associated with such a mean-value estimate?
- (4) How much maintainability improvement can reasonably be asked for?
- (5) Is there any margin for increasing the 20-minute specified value?

Again, the answers to these questions and the conclusions to be drawn depend on individual circumstances. To check for realism, the prediction technique developed in this study, those of MIL-STD-472, and others presented in the literature can be used as applicable.

Observed maintainability values of existing equipments obtained from several sources are presented in Tables XII through XV to provide historical data that can be used as a guide in assessing the realism of a specified value. The sources are identified in the tables according to the following numbered references:

1. RADC-TDR-63-85, Vol. 1, Maintainability Technique Study, Final Technical Report (Phase V), 5 February 1963. Prepared by RCA.
2. RADC-TN61-141, Maintainability Measurement and Prediction Methods for Air Force Ground Electronic Equipment (Phase III Progress Report), 15 June 1961. Prepared by RCA.
3. ARINC Research Corporation Publication 118-4-228, Maintainability of Shipboard Electronic Systems, 31 March 1961.
4. RADC-TR-68-398, Maintainability Prediction by Function, Final Report, August 1968. Prepared by Federal Electric Corporation. (The data from this source were accumulated at least in part from the AFM66-1 reporting system. Since this reporting system includes short-duration administrative delays in reporting man-hours, the data in Tables XII and XIII are also contaminated to some degree if they have been derived totally or in part from this source.)
5. Calculated from data accumulated by Federal Electric Corporation under Contract F30602-67-C-0194.
6. From data collected by ARINC Research Corporation under RADC Contract No. F30602-68-C-0047, Maintainability Prediction and Demonstration Techniques.

Table XII presents observed maintainability values for several classes of ground equipments. These data can be used to estimate the maintainability performance of a ground equipment if no details beyond the major functional classification are known.

Table XIII is an accumulation of maintainability data on a number of individual ground equipments arranged by common functional groupings. If the equipment under consideration can be

considered to be similar to one of the listed equipments, the maintainability values given in the table may be used in the absence of more precise estimating methods.

The values given in the table represent total active corrective maintenance, which generally includes preparation, fault-location, fault-correction, item-obtainment, checkout, and clean-up time, but excludes downtime due to administrative and logistic delays. In most cases, the results include maintenance actions for which no trouble was found. The occurrence of such events in a demonstration test must be considered in evaluating these data.

Tables XIV and XV are similar to Table XIII except that they represent data on airborne equipments monitored in this study. The observed maintainability values in Table XIV include maintenance events for which no trouble was found. Table XV does not include the "no trouble found" events and may therefore be more applicable for evaluating maintainability-demonstration index values.

TABLE XII

MAINTAINABILITY DATA FOR VARIOUS GROUND EQUIPMENT CLASSES

Equipment Class	$M_{CT}$	$\tilde{M}_{CT}$	$M_{0.95}$	Average MMH per Action	Data Source
Transceiver	0.88	0.36	3.2	1.1	4 and 5
Receiver	1.83	0.90	5.4	2.3	4 and 5
Transmitter	1.91	0.80	6.3	2.4	4 and 5
Display/Indicator	1.51	0.79	4.4	1.9	4 and 5
Data Processing	2.23	1.00	8.4	2.8	4 and 5
Frequency Power Supply	1.11	0.51	4.6	1.4	4 and 5
Identification Recognition	1.67	0.80	5.0	2.1	4 and 5
Multiplex	1.35	0.37	4.7	1.7	4 and 5
Exciter	1.75	0.90	5.4	2.2	4 and 5
Data Processor	1.19	0.33	3.7	1.5	4 and 5

TABLE XIII

MAINTAINABILITY DATA FOR VARIOUS  
GROUND EQUIPMENTS

Equipment	$M_{CT}$ (Hours)	$\tilde{M}_{CT}$ (Hours)	$M_{0.95}$ (Hours)	$\frac{MMH}{OH}$ $\times 10^3$	Average MMH per Action	Source
<u>Radio Transceivers</u>						
AN/MPH-14 (ARC-27)					1.0	4
AN/MPN-14 (GRA-54)					2.0	4
AN/GRC-131					0.2	4
AN/GRC-132					0.2	4
AN/GRC-113					0.3	4
AN/GRT-3 and GRR-7	1.05	0.80	2.8			1
AN/GKA-5	1.62				2.13	2
<u>Receivers</u>						
AN/FPS-16					4.4	4
AN/FPS-27	0.61	0.43	1.7		1.25	6
ARSR-1B					1.9	4
AN/FPS-30					1.2	4
AN/GSQ-74B					0.5	4
AN/FPN-47					2.8	4
AN/MPN-14					1.0	4
AN/FRC-102					1.9	4
AN/FRC-96					2.6	4
MW-503A					1.2	4
74A2					2.1	4
AN/GRC-66					3.0	4
AN/GRC-126					0.9	4
AN/TRN-17					3.0	4
AN/SRR-13A	2.34	1.20	7.0	1.86	3.6	3
AN/URR-35A	5.31	0.92		0.43	6.4	3
AN/SLR-2	2.50	1.60	6.3	16.41	4.12	3

(continued)

TABLE XIII (continued)

Equipment	$M_{CT}$ (Hours)	$\tilde{M}_{CT}$ (Hours)	$M_{0.95}$ (Hours)	$\frac{MMH}{OH} \times 10^3$	Average MMH per Action	Source
<u>Transmitters</u>						
AN/TPQ-18	1.29	0.81	4.2		15.1	4
AN/FPS-16					2.8	4
AN/FPS-27					2.9	6
ARSR-1B					0.8	4
AN/FPS-30					5.1	4
AN/GSQ-74B					1.1	4
AN/FPN-47					2.4	4
AN/MPN-14(T273B)					1.6	4
AN/MPN-14(ARC3)					1.3	4
AN/MPN-14(T867)					1.4	4
AN/FRC-102					3.0	4
AN/FRC-96					4.0	4
74A2					3.1	4
AN/GRC-66					2.2	4
AN/GRC-113					0.4	4
AN/FRT-37					1.9	4
AN/GRN-9C					4.5	4
AN/TRN-17					2.9	4
AN/MRN-13					1.8	4
AN/URN-5					2.1	4
AN/SRT-15	3.25	1.75	8.0	30.6	8.3	3
<u>Display Indicators</u>						
AN/TPQ-18					7.7	4
AN/FPS-16					1.9	4
AN/FPS-30					2.3	4
AN/GSQ-74B					1.3	4
AN/FPN-47					2.1	4
AN/MPN-14					1.5	4

TABLE XIII (continued)

Equipment	$\bar{M}_{CT}$ (Hours)	$\tilde{M}_{CT}$ (Hours)	$M_{0.95}$ (Hours)	$\frac{MMH}{OH}$ $\times 10^3$	Average MMH per Action	Source
<u>Display Indicators</u>						
AN/MPN14(MF49)					1.4	4
AN/MPN14(MPA31)					1.3	4
74A2					1.3	4
AN/GSW-5					0.4	4
AN/GSA-51	0.90	0.53	3.2		1.9	6
AN/UPA-35					4.6	4
AN/FSA-14					1.9	4
AN/FSA-26					1.9	4
AN/SPA-4	2.50	1.00	12.0	8.57	5.21	3
AN/SPA-4A	4.90	0.78		16.67	11.89	3
AN/3PA-8A	1.96	0.65	8.0	6.17	3.34	3
AN/SPA-8C	4.20	1.47		7.86	5.18	3
VL-1	1.20	0.43		0.83	2.33	3
ECCM(FPS-27)	1.16	0.94	3.1		1.57	6
<u>Data Processing</u>						
AN/TPQ-18					5.1	4
AN/FPS-16					2.5	4
AN/FPN-47					1.6	4
AN/FST-2	0.93				1.27	2
AN/GSW-5					0.6	4
AN/GSW-10					0.4	4
AN/GSA-51	0.70	0.33	2.3		1.50	6
Computer (no mili- tary nomenclature)					1.8	4
Computer (no mili- tary nomenclature)					2.3	4
Recorder/Reproducer (GSA-51)	2.03	1.27	4.1		3.79	6

TABLE XIII (continued)

Equipment	$\bar{M}_{CT}$ (Hours)	$\tilde{M}_{CT}$ (Hours)	$M_{0.95}$ (Hours)	$\frac{MMH}{OH_3}$ $\times 10^3$	Average MMH per Action	Source
<u>Data Processing</u>						
Punch Card System (GSA-51)	.73	.605	2.0		1.35	6
<u>Radar</u>						
AN/FPS-6	1.57	1.25	3.5			1
AN/FPS-20	1.11				1.58	2
AN/SPS-8A	3.50	1.26	10.5	28.28	5.51	3
AN/SPS-2C	2.80	1.03		8.04	3.82	3
AN/SPS-12	2.50	1.50	7.3	22.58	4.17	3
AN/SPN-8	1.90	1.43	7.2	15.21	2.54	3
AN/SPN-12	2.90	2.00		18.91	4.88	3
<u>Navigation</u>						
AN/URN-3	9.60	3.20		55.29	21.85	3
AN/URN-4	7.20	0.50		20.26	14.98	3
AN/SPN-7A	2.20	1.80		7.62	2.90	3
<u>Frequency/Power Supply</u>						
AN/FPS-30					1.3	4
AN/MPN-14					1.3	4
74A2					2.0	4
AN/TRN-17					0.9	4
<u>Identification/ Recognition</u>						
AN/UPX-6					2.3	4
KY-274					2.1	4

TABLE XIII (concluded)

Equipment	$M_{CT}$ (Hours)	$\tilde{M}_{CT}$ (Hours)	$M_{0.95}$ (Hours)	$\frac{MMH}{OH} \times 10^3$	Average MMH per Action	Source
<u>Identification/ Recognition</u>						
KY-248					1.3	4
AN/GPX-8A					1.7	4
AN/MPN14(MPA24)					1.5	4
AN/UPX-1A	2.90	1.30	18.0	8.10	4.50	3
<u>Multiplex</u>						
AN/FCC-22					1.9	4
AN/FCC-32					1.9	4
TCS-600					0.9	4
AN/FGC-5					2.7	4
AN/FGC-29					1.8	4
AN/FGC-61					1.7	4
<u>Exciter</u>						
AN/FRC-102					2.2	4
AN/FRC-96					3.2	4
AN/GRC-66					1.4	4
AN/GRC-113					0.3	4

TABLE XIV

MAINTAINABILITY DATA FOR AVIONIC EQUIPMENT  
(NO-TROUBLE-FOUND ACTIONS INCLUDED)

Equipment	$M_{CT}$ (Hours)	$\tilde{M}_{CT}$ (Hours)	$M_{0.95}$	Average MMH per Action	Source
<u>Navigation/Radio Receiver</u>					
APN-151	0.99	0.92	1.80	1.46	6
ADF-73	0.26	0.19	0.74	0.26	6
<u>Electromechanical Navigation Computer</u>					
ASN-35	0.58	0.52	1.24	0.81	6
ASN-24	0.70	0.49	2.29	0.96	6
<u>Radio Communication</u>					
ARC-109	1.17	1.00	2.94	2.73	6
ARC-90	0.51	0.35	1.79	0.82	6
<u>High-Power Radar</u>					
APQ-110	1.03	0.70	2.88	1.81	6
AJQ-20	1.77	1.47	4.02	3.11	6
APQ-113	1.20	0.92	4.04	3.04	6
APS-109	1.75	1.32	5.79	4.60	6
APN-59	0.94	0.59	3.29	1.54	6
<u>Low-Power Navigational and IFF Transmitters and Receivers</u>					
APN-167	0.64	0.52	1.82	1.31	6
APN-147	0.81	0.54	2.72	1.45	6
APN-21	0.71	0.50	2.05	1.15	6

TABLE XV  
MAINTAINABILITY DATA FOR AVIONIC EQUIPMENT  
(NO-TROUBLE-FOUND ACTIONS EXCLUDED)

Equipment	$M_{CT}$ (Hours)	$\tilde{M}_{CT}$ (Hours)	$M_{0.95}$	Average MMH per Action	Source
<u>Navigation/Radio Receiver</u>					
APN-151	1.30	0.95	2.01	1.63	6
ADF-73	0.34	0.27	0.96	0.34	6
<u>Electromechanical Navigation Computer</u>					
ASN-35	0.68	0.63	1.28	1.23	6
ASN-24	0.86	0.64	1.96	1.19	6
<u>Radio Communication</u>					
ARC-109	1.24	1.07	3.10	2.91	6
<u>High-Power Radar</u>					
APQ-110	1.22	0.99	3.65	2.35	6
AJQ-20	2.46	2.20	4.87	4.71	6
AFQ-113	1.40	0.90	7.18	3.34	6
APN-59	1.02	0.65	3.46	1.66	6
<u>Low-Power Navigational and IFF Transmitters and Receivers</u>					
ARN-21	0.82	0.63	2.24	1.35	6

### 5.3.4 Applicability of Requirements to the Demonstration Environment

In the discussion of the results of the maintainability demonstration survey in Section II, it was noted that the most frequently cited difficulty was the difference between test environment and field environment. In an RADC study<sup>3</sup>, a comparison of demonstration-test results with field operational results for seven systems revealed wide discrepancies. The operational field MTR was always greater. Although the field data may have been contaminated with some undesirable factors such as administrative-time delays, the observed differences are still quite illuminating.

It is apparent that the closer the test environment to the expected field environment, the more meaningful the demonstration test, and that every effort should be made to achieve such similarity. Specific reasons for biases due to test environment are outlined in this section.

Unless a Category III type test is to be performed, demonstration environments will differ in some respects from the field environment. Because such differences do exist, a maintainability demonstration requirement based on operational goals should not be applied unless its applicability to the demonstration conditions is first considered.

As a general principle, the specified value based on operational goals and conditions must be suitably adjusted to reflect the maintenance environment governing the demonstration. Often, this is a difficult principle to adhere to. With an avionics equipment, for example, a certain amount of time will be spent in the field just reaching the equipment in the aircraft, and the time to locate the malfunction and complete repairs and checkout is a function of this accessibility factor. If the demonstration test is not to take place in the aircraft (and this is often the case) there is the question of whether the specified value should be adjusted and how much.

It might be possible to construct a mockup to simulate the actual conditions, thus eliminating the need for adjustment. Generally, this type of simulation will not be possible, and field and test conditions must be carefully analyzed and their effects quantitatively assessed. Table XVI lists various factors to be considered in evaluating the applicability of a specified maintainability index. Table XVII lists some specific causes of discrepancies that are classified as yielding either pessimistic or optimistic results.

---

<sup>3</sup>A. Coppola and J. Deveau, "Reliability and Maintainability Case Histories", Annals of Reliability and Maintainability, Vol. 6, 1967, pp. 582-586.

**TABLE XVI**

**FACTORS AFFECTING THE SUITABILITY OF A SPECIFIED MAINTAINABILITY  
INDEX FOR MAINTAINABILITY DEMONSTRATION**

**Physical Equipment**

Stage of completion  
Similarity to production items  
Physical location  
Interfacing equipment

**Test Location and Facility**

Lighting factors  
Weather factors  
Space factors

**Test Team**

Organization  
Training and experience  
Indoctrination

**Support Items**

Tools  
General and special test equipment  
Spares availability  
Technical manuals

**Operational Factors**

Mode of equipment operation  
Procedures for instituting maintenance  
Procedures for fault selection

TABLE XVII

CAUSES OF DISCREPANCIES BETWEEN TEST AND FIELD RESULTS

A. Causes of Optimistic Test Results

1. The demonstration maintenance technicians are not representative of typical Air Force personnel because they have more education and training or greater knowledge of the equipment design.
2. The monitoring situation imparts to the technician an urgency not normally encountered in the field.
3. Known probable tasks are rehearsed beforehand.
4. Necessary support equipment is readily available.
5. Observed times are not contaminated with such factors as administrative or logistic delay, as field results sometimes are.
6. Difficult-to-isolate faults such as intermittencies and degradation failures are not simulated.

B. Causes of Pessimistic Test Results

1. The technicians are not familiar with the equipment and have not acquired the necessary experience for rapid fault isolation.
2. Field and procedural modifications to reduce maintenance time have not yet been made.
3. Initial manuals may be incomplete or require revision.
4. The monitoring situation can adversely affect the technician's performance.

## 5.4 RISK ASSIGNMENT

### 5.4.1 General

There are generally two risks involved in a demonstration test:

- (1) Producer's risk,  $\alpha$  -- the probability of rejection if the maintainability characteristic is at the desired level
- (2) Consumer's risk,  $\beta$  -- the probability of acceptance if the maintainability characteristic is at the minimum acceptable (or undesirable) level

Ideally,  $\alpha$  and  $\beta$  would be equal to zero; granting that this is impossible, very small values of  $\alpha$  and  $\beta$  -- on the order of 0.001 -- are desirable. Such small values are also impractical since, as discussed in Section III, the selection of  $\alpha$  and  $\beta$  associated with the  $H_0$  and  $H_1$  values for maintainability dictates the sample size. For  $\alpha$  and  $\beta$  on the order of 0.001, sample sizes far exceeding available test resources will usually be required.

For example, consider a test of the mean of a lognormal distribution such as the following:

$$H_0: \mu = \mu_0 = 30 \text{ minutes}$$

$$H_1: \mu = \mu_1 = 45 \text{ minutes}$$

As shown in Section VII (Test Number 1), the necessary sample size for this test is given by the equation

$$n = \frac{(Z_\alpha \mu_0 + Z_\beta \mu_1)^2}{(\mu_1 - \mu_0)^2} (e^{\sigma^2} - 1)$$

where  $Z_\alpha$  and  $Z_\beta$  are the normal deviates corresponding to the  $(1 - \alpha)^{\text{th}}$  and  $(1 - \beta)^{\text{th}}$  percentile of a normal (0, 1) distribution and  $\sigma^2$  is the variance of the logarithm of maintenance time. If  $\alpha = \beta$  and  $\sigma^2 = 1$  are assumed, then

$$n = \frac{Z_\alpha^2 (30 + 45)^2}{(45 - 30)^2} (e^1 - 1) = 43Z_\alpha^2$$

Figure 6 shows the relationship between  $n$  and the risk values. In the figure, it can be seen that if  $\alpha = \beta = 0.10$ , 70 observations (a reasonable number) are required. If  $\alpha$  and  $\beta$  are reduced to 0.01, about 230 observations are necessary, and for  $\alpha = \beta = 0.001$ , a sample size of more than 400 is called for.

Most development budgets and schedules will not allow for a test requiring 400 sample observations even if the observations are to be simulated. In fact, even a sample size of 70 may tax available resources, and for this illustrative case, risks on the order of 0.15 or 0.20 may be necessary.

It is not necessary, of course, for  $\alpha$  to equal  $\beta$ . If, for example, the need for the equipment is great and a 45-minute mean time to repair can be tolerated (perhaps with later improvement by modification and appropriate training, manning, and support planning), the  $\beta$  risk may be set at a higher level, say 0.25. This means that there is a relatively low risk of rejecting good equipment and a higher risk of accepting a minimum acceptable equipment.

#### 5.4.2 Use of Prior Information in Risk Trade-off

The choice of  $\alpha$  and  $\beta$  is also one involving trade-offs. From a decision-theory viewpoint, the trade-off can be normalized to a cost criterion based on the following:

- (1) Cost of testing (sample size)
- (2) Cost of rejecting good equipment
- (3) Cost of accepting poor equipment

While (1) can generally be costed in terms of manpower, facilities, and time, (2) and (3) are more difficult to assess quantitatively. Assuming that prior information is available for estimating at least relative values associated with the three costs, two simplified approaches employing decision-theory concepts for selecting  $\alpha$  and  $\beta$  are discussed below. For convenience, the maintainability characteristic of interest will be denoted by  $M$ , and specified  $H_0$  and  $H_1$  values by  $M_0$  and  $M_1$ , respectively. Also let

$C_0$  = Cost of rejection if  $M = M_0$

$C_1$  = Cost of acceptance if  $M = M_1$

##### 5.4.2.1 Minimax Criterion

The minimax criterion is used when it is desirable to avoid extremely high costs. In order to use this criterion,

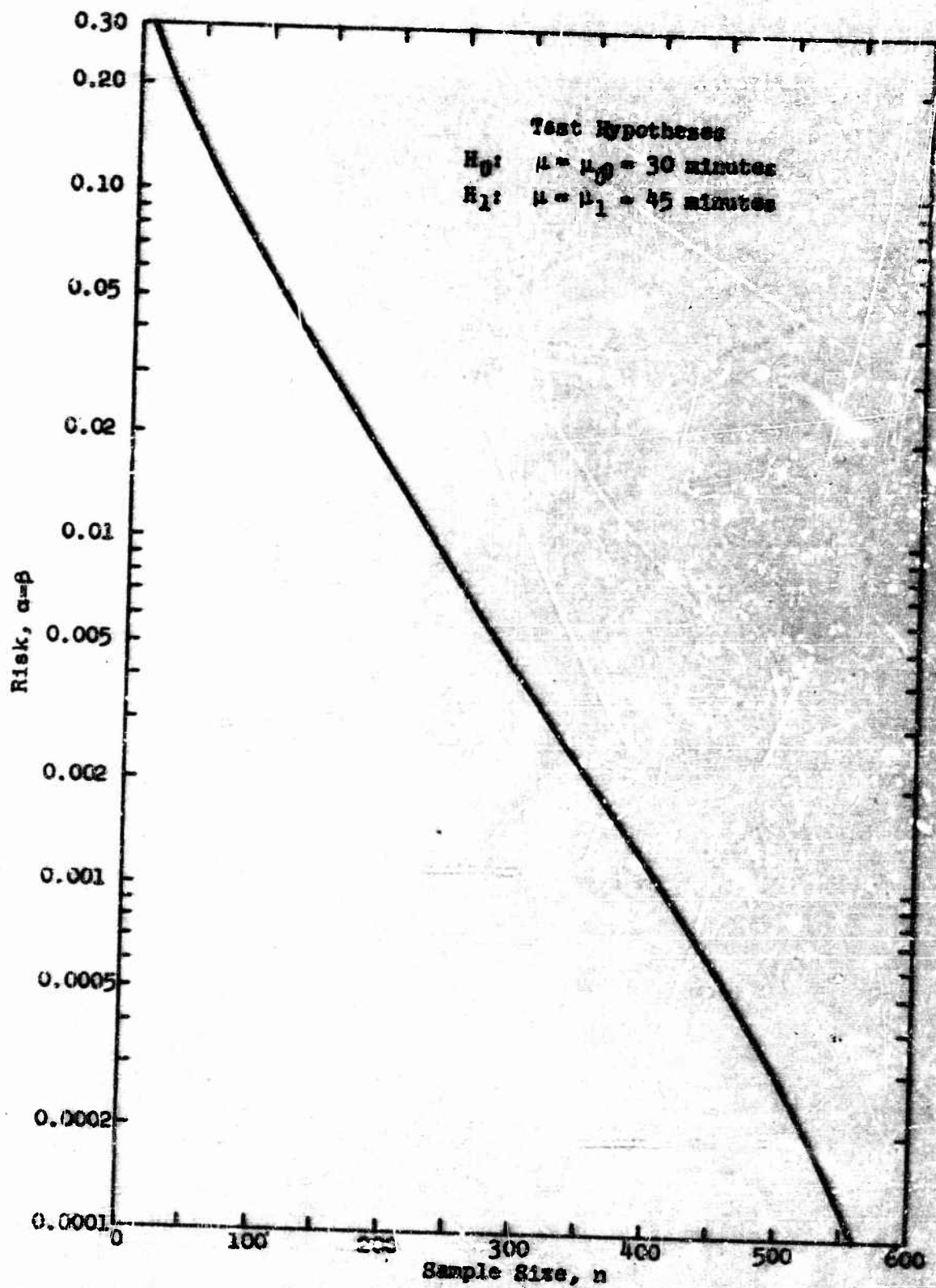


FIGURE 6

SAMPLE-SIZE REQUIREMENT FOR TEST OF  
 1) NORMAL MEAN AS FUNCTION OF RISK ( $\alpha = \beta$ )

for a given combination of  $\alpha$  and  $\beta$ , say  $(\alpha_1, \beta_j)$ , compute the following:<sup>4</sup>

- (1)  $L_{1j}(M_0) = C_0\alpha_1 + C_{1j}(M_0)$
- (2)  $L_{1j}(M_1) = C_1\beta_j + C_{1j}(M_1)$
- (3)  $L_{1j} = \text{Max} [L_{1j}(M_0), L_{1j}(M_1)]$

where

$C_{1j}(M_k)$  = Test costs associated with  $(\alpha_1, \beta_j)$   
if  $M = M_k$  ( $k = 0$  or  $1$ )

$L_{1j}(M_k)$  = Total cost if  $M = M_k$  ( $k = 0, 1$ ) and  $\alpha = \alpha_1, \beta = \beta_j$

$L_{1j}$  = Maximum cost if  $\alpha = \alpha_1, \beta = \beta_j$

Generally  $C_{1j}(M_k)$  will be a function of the sample-size requirements dictated by the  $\alpha_1, \beta_j$  pair and will not depend on  $M$  except for sequential tests,<sup>1</sup> for which the average value of  $n$  given  $M = M_k$  can be used.

The  $\alpha, \beta$  risk pair to select is that which has the minimum value of  $L_{1j}$ . By this criterion the selected risks are such that the maximum possible costs are minimized.

**Example:** Consider the illustrative test discussed above. For simplicity, assume that the values of  $\alpha$  and  $\beta$  to be considered are restricted to 0.05, 0.10, 0.20. Some possible risk pairs and associated sample sizes, from the previous equation, are as follows:

Pair (i,j)	$\alpha$	$\beta$	$n_{1j}$
11	0.05	0.05	116
12	0.05	0.10	87
13	0.05	0.20	58
21	0.10	0.05	96
22	0.10	0.10	70
23	0.10	0.20	44
31	0.20	0.05	75
32	0.20	0.10	52
33	0.20	0.20	30

<sup>4</sup> These equations are based on the assumption that no costs are associated with an accept decision if  $M=M_0$ , or reject decision if  $M = M_1$ , except for the test costs.

Cost considerations lead to the following relationships:

$$C_0 = \$50,000$$

$$C_1 = \$40,000$$

$$C_{1j} = \$2000 + n_{1j}^2$$

The results of the necessary computations are shown in Table XVIII. For each pair, the maximum value of  $L_{1j}$  is underscored. The minimum of these maximum values is seen to be \$11,900, which is yielded by the pair  $\alpha = 0.10$ ,  $\beta = 0.10$ .

TABLE XVIII  
COMPUTATIONS FOR OBTAINING MINIMAX  
RISKS FOR ILLUSTRATIVE EXAMPLE

Index		Risks		Costs	
1	j	$\alpha$	$\beta$	$L_{1j}(M_0)$	$L_{1j}(M_1)$
1	1	0.05	0.05	<u>\$17,956</u>	\$17,456
1	2	0.05	0.10	<u>12,069</u>	<u>13,569</u>
1	3	0.05	0.20	<u>7,864</u>	<u>13,364</u>
2	1	0.10	0.05	<u>16,216</u>	<u>13,216</u>
2	2	0.10	0.10	<u>11,900*</u>	10,900
2	3	0.10	0.20	<u>8,936</u>	<u>11,936</u>
3	1	0.20	0.05	<u>17,625</u>	<u>9,625</u>
3	2	0.20	0.10	<u>14,704</u>	<u>8,704</u>
3	3	0.20	0.10	<u>12,900</u>	10,900

\*Minimum of maximum values.

#### 5.4.2.2 Bayes Strategy

For the Bayes approach, prior information or subjective evaluation is required to estimate the following:

$$P_0 = \text{probability } M = M_0$$

$$P_1 = 1 - P_0 = \text{probability } M = M_1$$

Then for each pair (1,j) the expected cost is computed:

$$E_{1j} = P_0 [C_0 \alpha_1 + C_{1j}(M_0)] + P_1 [C_1 \beta_j + C_{1j}(M_1)]$$

The pair for which  $E_{1j}$  is a minimum is selected.

In this procedure, the risks are selected to minimize the expected costs.

**Example:** Assume that it can be reasonably estimated from past performance data, in conjunction with evaluation of the maintainability program efforts, that  $P_0 = 0.70$ ,  $P_1 = 0.30$ . The values associated with this prior distribution are as follows.

Index		Risks		Expected costs, $E_{1j}$
i	j	$\alpha$	$\beta$	
1	1	0.05	0.05	\$17,806
1	2	0.05	0.10	12,519
1	3	0.05	0.20	9,514*
2	1	0.10	0.05	15,316
2	2	0.10	0.10	11,600
2	3	0.10	0.20	9,836
3	1	0.20	0.05	15,225
3	2	0.20	0.10	12,904
3	3	0.20	0.20	12,300
*Minimum Value.				

From the above listing, it is seen that the risk  $\alpha = 0.05$ ,  $\beta = 0.20$  minimizes expected cost. If the prior probabilities were  $P_0 = P_1 = 0.50$ , the pair  $\alpha = 0.10$   $\beta = 0.20$  would be optimal. With the prior estimates of  $P_0$  and  $P_1$ , the expected cost without testing can also be evaluated. If no testing is performed and the equipment is to be accepted upon delivery, the expected cost is simply

$$(P_1)(C_1) = (0.30)(40,000) = \$12,000.$$

For this example, the decision not to test is unwise. However, where testing is quite costly and past performance indicates a high probability of a satisfactory product, this type of evaluation might indicate that, from the viewpoint of economy, little or no testing is the preferred choice.

#### 5.4.2.3 Summary of Decision-Theory Approach

The two decision-theory approaches described above might be criticized on the basis that only the  $H_0$  and  $H_1$  values for  $M$  are considered. More extensive procedures can be used, but they require prior information and cost relationships that are not generally available.

In defense of the procedure, it can be said that in conventional sampling procedures, in which  $\alpha$  and  $\beta$  are more or less arbitrarily chosen, two levels of maintainability are also considered. Moreover, the  $M_0$  and  $M_1$  values and their associated

risks do determine the complete O.C. curve. Choosing  $\alpha$  and  $\beta$  from a decision-theory viewpoint does provide some cost control for the test procedure and thus has economic advantage over non decision-theory approaches.

## SECTION VI

### SELECTION OF MAINTENANCE-TASK SAMPLE

#### 6.1 GENERAL

There are two basic approaches for selecting a sample of maintenance tasks for the demonstration:

- (1) Observe maintenance tasks as they occur naturally in an operational or simulated operational situation
- (2) Induce faults in the system and observe the maintenance actions to correct these faults

The terms "natural failures" and "fault inducement" will be used to distinguish these two approaches.

For the fault-inducement approach, a decision must be made on the type of sampling procedure to be used. The usual choice is between stratified sampling and simple random sampling.

In this section guidelines are offered for evaluating the applicability of the two basic approaches, obtaining maintenance-task samples, and choosing the appropriate sampling design and procedure.

#### 6.2 NATURAL VERSUS INDUCED FAILURES

In most cases, the choice of natural or induced failures must be made early in the development program since the natural-failure approach can be used only if the program schedule allows enough time to obtain the required number of maintenance tasks. This allowable time is, of course, related to reliability.

If  $\theta$  is the MTBF of an equipment, the average number of operating hours that will be required to yield  $n$  failure occurrences is  $n\theta$ . For equipments with MTBFs of hundreds of hours and required sample sizes of 30 to 70, the number of required equipment operating hours can easily exceed 10,000 (e.g., 50 samples from an equipment with an MTBF of 200 hours). If three such equipments are available for test and the equipments are operated 16 hours a day, an average of more than 200 days would be required to complete the demonstration.

Because of time requirements of this magnitude, most maintainability demonstrations are based on the fault-inducement approach, by which the demonstration can be completed in a few days.

With respect to realism and applicability, the natural-failure approach is clearly the preferred choice. The major disadvantage of inducing faults is that there is no guarantee that these faults are representative of those which will be seen in operation. In addition, the fault-inducement plan provides information for rehearsal, which will naturally bias the test in a direction that is favorable to the contractor. Unfortunately, there have been indications that such rehearsals have taken place in demonstration tests.

Because of the problems associated with fault inducement, the following general recommendations are made:

- If the schedule can allow for natural failures, then this type of sampling procedure is preferred. Category III type tests will logically fall within this class.
- If the complete demonstration cannot be completed with only naturally occurring failures, a combination of the two approaches should be used. One possibility is to take advantage of the reliability-demonstration test and include in the sample the maintenance times needed to correct faults that occurred in the reliability demonstration. Close coordination between the two test groups will be required.
- If a natural-failure test cannot be conducted, any natural failures that do occur during the induced-failure test should be included in the sample.

### 6.3 FAULT-INDUCEMENT PROCEDURES

The first criterion for judging the suitability of a fault-inducement procedure is whether it leads to a series of maintenance tasks that are representative at the level of maintenance specified. Thus the maintenance tasks generated by fault inducement should be representative with respect to the following:

- Engineering and maintenance factors such as symptom indications and required repair procedures
- Frequency of occurrence

The latter will be considered in Subsection 6.4, which describes the sampling procedure.

Frequently used methods for inducing a fault are to tape a connector pin or disconnect a lead to simulate an open, grounding a wire or pin to simulate a short, inserting a known faulty

part or blown fuse, and removing a circuit card or wire.

The method used should not, of course, provide the technician with information that he would not normally receive under actual maintenance conditions. Taping a pin, therefore, would not be an acceptable method if the level of maintenance was such that the technician could easily spot this type of simulated fault. If, however, maintenance were at the module level and the pins were internal, such a method would be acceptable.

More than avoiding obvious simulations is required for adequate fault inducement. A review of past maintainability demonstrations indicates that the fault-inducement methods involve a great number of disconnects, card removals, wire groundings, and the like to simulate either shorts or opens with little or no regard to other type failures.

These methods are relatively easy to accomplish; they can be controlled so that the equipment is not damaged and the induced fault can be easily corrected upon completion of the maintenance observation. However, they may not lead to a representative set of maintenance tasks. Failures resulting from out-of-tolerance or degradation conditions or intermittencies and those of a secondary nature leading to a multiple failure occurrence usually cannot be simulated by these simple methods.

There are several possible approaches for inducing non-catastrophic failures:

- Replacement of a good part, circuit, or assembly with an identical item with an appropriate type failure
- Insertion of extra nondetectable parts such as a bypass resistor to simulate an out-of-tolerance condition
- Deliberate misalignment
- Use of cold-solder joints to induce intermittencies

Consideration should be given to including failures that cannot be attributed to piece-part inherent reliability such as nicked insulation, broken wires, and items abused by operation or through maintenance (e.g., bent pins). Secondary failures, which are a result of a primary failure, must be considered if their frequency of occurrence is not negligible.

The strongest argument for avoiding noncatastrophic failures is that they are difficult to induce by simulation, and this cannot be denied. However, the cost of such failure inducement should be small relative to the total cost of the demonstration. More important, the additional investment for including more than just open- and short-type failures provides

a more representative set of tasks and, therefore, eliminates a possibly serious biasing factor

One approach to achieving this type of noncatastrophic representation is for the contractor to retain parts, circuit cards, assemblies, etc., that have been rejected during development, reliability, and quality-control tests for possible use in the maintainability demonstration. This is particularly important for failures that are difficult to simulate, such as intermittencies and instability.

From the foregoing discussion, it is apparent that the types of faults to be simulated and the methods of simulation must be considered early in the development program. Such early planning will make it possible to use the information from development-type tests to achieve realism through effective simulation. Because of the intricacies that may be involved in simulating certain types of malfunctions, the guidance of design-engineering and reliability personnel will generally be required in planning the fault-inducement procedures. The results of failure-mode and effects analyses (FMEA), reliability predictions, and reliability tests, are particularly applicable.

#### 6.4 DESIGN OF MAINTENANCE-TASK SAMPLE

Randomness, lack of bias, representativeness, and efficiency are several of the criteria for evaluating a sample design. In an experiment such as testing the efficacy of a drug on laboratory animals or polling a relatively known and stable population on a political question, the use of such criteria is possible, meaningful, and prudent. In maintainability demonstration, however, the population of interest, maintenance-task times, does not usually exist at the time of the test, and the anomaly of sampling from a nonexistent population certainly complicates sample design.

Sampling from a population whose specific characteristics have not been previously observed is not unique, however. For example, it is possible to estimate the number of different types of fish in a newly discovered lake from a sample designed by stratifying the lake by areas, if general information is available on the habitats of different fish. Care must be taken to eliminate biases in methods used for catching the fish, time of day, season, etc. -- that is, the procedures used should not result in obtaining one species of fish in a greater proportion than actually exists in the population.

Similarly, the basic objective in obtaining samples from naturally occurring failures or from fault-inducement procedures should be to yield unbiased estimates of the maintenance parameter of interest. It is probably impossible to avoid bias

completely. For example, the inducement of a fault that can seriously damage the equipment is almost always prohibited. However, possibly serious biasing factors, such as inducing only easily repaired faults, should be closely controlled. If the direction and magnitude of unavoidable remaining biasing factors is known, appropriate adjustments to the specified values may have to be made, as discussed previously.

#### 6.4.1 Sample Design for a Natural-Failures Test

For demonstration based on natural failures, there is little flexibility in sample selection. As failures occur or preventive maintenance actions are called for, observers record the appropriate maintainability characteristic when the maintenance action is initiated. The small number of available equipments and test environments and limited test time will generally preclude the use of sophisticated sampling procedures and require following an "observe everything we can" philosophy.

Those responsible for sample design should review the conditions under which the natural failures will occur. If the analysis indicates that certain types of actions will be unlikely, as might be the case if the operational mode under test does not require using a portion of the system, then some fault inducement may be necessary. Therefore, "filling out the sample" should be considered for a natural-failures test.

To aid further in determining representativeness, it is recommended that a list of the various possible maintenance tasks be developed, possibly grouped by similarity of required procedures and expected task times. The frequency of occurrence of each group must then be estimated on the basis of such factors as number of items in the system that can lead to performance of the particular task or group of tasks, the reliability of these items, and the operating duty cycle.

Such a list will then provide an estimated relative frequency of occurrence of various tasks, which can be used to evaluate possible biases in the observed sample. Since the list must be developed partly on the basis of estimates, such as failure rate, there is no guarantee that the checklist is "right" and the sample "wrong". However, such a list can provide a warning signal for large discrepancies, which should then be investigated in greater detail.

Since the development of such a list is the basic approach to designing a fault-inducement sample, the details are presented in the following subsection, which deals with that method of sample selection.

## 6.4.2 Sample Design for a Fault-Inducement Test

### 6.4.2.1 Simple Random Versus Stratified Sampling

The basic choice in designing a sampling procedure for inducing faults is between simple random sampling and stratified sampling.

A simple random sample is one in which all possible samples of  $n$  units out of the population have an equal chance of being chosen. A stratified random sample is one in which the total population is divided into subpopulations or strata and sample sizes for each stratum are then determined according to selected criteria. Random sampling is then performed within each subpopulation.

Because there is no physical population as such from which to sample, it is necessary first to develop a hypothetical population of maintenance tasks. This hypothetical population provides the basis for sampling by the fault-inducement procedure for both random and stratified sampling.

A simplified scheme for presenting this hypothetical population is shown in Table XIX. The maintenance-task groups represent all the different types of maintenance tasks that may be performed, ranging from simple adjustments to complicated mechanical repairs. Similar tasks are usually grouped together. The expected number of task occurrences within a maintenance-task group can be estimated by the equation

$$E_i = \sum_j \lambda_{ij} k_{ij} T$$

where

$\lambda_{ij}$  is the failure rate of the  $j^{\text{th}}$  item in the  $i^{\text{th}}$  group

$k_{ij}$  is the duty-cycle factor for the  $j^{\text{th}}$  item in the  $i^{\text{th}}$  group ( $0 < k_{ij} \leq 1.0$ )

$T$  is the average mission time

The sum of the  $E_i$ 's is the expected number of maintenance actions in  $T$  hours, and this is used to obtain the relative task-occurrence probabilities,  $p_i$ . Details of developing such a table are presented in Subsections 6.4.2.2 and 6.4.2.3. Table XIX will be used here to discuss the distinction between simple and stratified sampling.

**TABLE XIX**  
**MAINTENANCE-TASK POPULATION**

Maintenance-task Group	Expected Number of Occurrences in T Hours	Relative Maintenance-Task Population
1	$E_1$	$E_1/E = p_1$
2	$E_2$	$E_2/E = p_2$
$\vdots$	$\vdots$	$\vdots$
1	$E_1$	$E_1/E = p_1$
$\vdots$	$\vdots$	$\vdots$
s	$E_s$	$E_s/E = p_s$
	$E = \sum_{i=1}^s E_i$	$\sum_{i=1}^s p_i = 1.0$

To select a simple random sample of size  $n$ ,  $n$  random numbers between 0 and 1 can first be drawn, such as from a random-number table. If a selected random number,  $x$ , is in the interval  $0 < x \leq p_1$ , a fault generating a task in the first group is induced. If it is in the interval  $p_1 < x \leq (p_1 + p_2)$ , the second type of task is generated.  $[(p_1 + p_2) < x \leq (p_1 + p_2 + p_3)]$  defines the interval for generating the third type of task, and so forth. By this procedure, all possible samples of size  $n$  have an equal chance of being observed.

The most commonly used method of stratified sampling in maintainability demonstration is proportional stratified sampling. In this method, the sample size from each stratum (e.g., maintenance-task group) is proportional to the population size of the stratum. Thus if there were five strata with relative population sizes of 5, 20, 20, 25, 30, and a total sample of 50 observations were to be made, 2 or 3 observations would be selected from the first, or smallest, stratum, 5 from the second and third, 10 from the fourth, 12 or 13 from the next, and 15 from the largest stratum.

It is noted that for a simple random sample, it is unlikely that such a "natural" sample-size distribution would occur, but as the total sample size increases, the simple random and proportional stratified samples will tend to coincide. For a small sample, however, it is quite possible that a sample from a specific stratum will not be observed in a random sample. For example, if the above relative population sizes are used, with a total sample size of 20, there is a  $(0.95)^{20} \approx 0.112$  probability that no samples from the smallest stratum will be observed, while a proportional sample would dictate that one such observation be drawn.

The most common applications of the theory of hypothesis testing are based on the assumption of simple random sampling. On the other hand, stratified sampling is often employed in large-scale survey work to take advantage of the greater precision that such a procedure offers in obtaining estimates of population parameters.

Historically, proportional stratified sampling has been used in maintainability demonstration (MIL-STD-471, Appendix A, Task Selection Method), probably because identification of the maintenance-task population requires a form of stratification and because stratified sampling can ensure that at least one sample observation from each selected stratum will be included in the sample, thus providing some psychological assurance of representativeness.

There are several advantages to stratification. In some cases, a hypothesis concerning one portion of the system or one type of maintenance may be of interest and, therefore, there may be a sample-size requirement for each of the appropriate subpopulations. For example, if a maintainability test is concerned with total maintenance man-hours, it may be advisable to treat corrective- and preventive-maintenance actions as two subpopulations so that inferences can be made for each as well as for the total population of maintenance actions.

A second advantage is that there may be cases in which administrative considerations will dictate the use of stratification. For example, in conducting a maintainability-demonstration test, it may be advisable to consider the electronic and mechanical portions separately, and the sampling approach in each may differ, thus necessitating a stratified approach.

The third major advantage of stratification is statistical. By stratifying a heterogeneous population into homogeneous strata, the variation within each stratum is minimized so that an estimate of each stratum mean can be obtained with a relatively small sample. These strata means are then combined so

that the overall mean estimate has smaller variance than that of simple random samples. This concept can be extended to obtain a stratified plan that minimizes testing costs where costs of testing vary over the strata elements.

The major disadvantage of stratification is that while it is an effective method for increasing the precision of sample estimates, it introduces considerable complexity in the necessary analytical procedures for hypothesis testing. For example, if the hypothesis under test concerns the mean of a normal distribution, the well known "t" test is usually applied under simple random sampling. When stratified sampling is used, such simple application may be impossible since even though the overall population is normally distributed, the distribution within each stratum may be far from normal. Most textbook discussions of confidence-interval and hypothesis-test inference are therefore based on the assumption of simple random sampling.

The procedures currently used to analyze the results of maintainability-demonstration tests employing proportional stratified sampling are based on the assumption that only simple random sampling has been performed.

As mentioned above, for large sample sizes a proportional stratified sample and a simple random sample will yield essentially similar results. Therefore, for large sample sizes (at least 50 observations) and a small number of strata (e.g., no more than 10), the use of simple random-sampling analysis procedures for a proportional stratified sampling can be considered acceptable if one is willing to forego the greater precision generally offered by stratified sampling.

It should be noted, however, that for a stratified random sample, while all elements in the population may have equal chances of appearing in the sample, the sample observations are not independent with regard to order. For if  $(n-1)$  observations have been made, the last item must be selected from a particular stratum, whose identity depends on the previously obtained sample observations. This type of dependence violates one of the principal assumptions of standard sequential-sampling plans; therefore, a stratified random sample is generally not appropriate for sequential testing as commonly used.

To summarize, stratified sampling will yield more efficient tests than simple random sampling provided the following conditions are satisfied:

- There is a good basis for stratification.
- The variance within each stratum is small.

- The strata population sizes are known.
- Appropriate analytical procedures are available.

If simple random-sampling analytical procedures are employed, the main advantage of sampling by proportional stratification is psychological in that assurance is provided for including samples from each of the stratified populations. It is believed that this is the major justification for stratified sampling in maintainability demonstration.

Table XX provides guidelines for using simple random and stratified sampling by presenting a summary comparison for several major factors.

#### 6.4.2.2 Stratification by Maintenance-Task Groups

As indicated above, regardless of whether simple random or stratified sampling is used, a hypothetical maintenance-task population must first be developed when faults are being induced. Identifying the maintenance-task groups is the first step towards this goal. For discussion purposes, the terms maintenance-task group and maintenance-task stratum are used synonymously.

For practical reasons, stratification or grouping is usually performed for both simple random sampling and stratified random sampling. The difference between the two is that for stratified sampling, the number of samples to be taken from each stratum is predetermined, while for simple random sampling, the number of samples to be taken from each stratum is a random variable.

For discussion purposes, comments will be restricted to stratified sampling, since they will also generally apply to simple random sampling when task selection by fault inducement is being considered.

The first task in stratification is choosing criteria by which to stratify. This involves the characteristic by which to stratify, the number of strata, and the boundaries defining the individual strata.

For a large-scale survey using stratified sampling and appropriate analytic procedures, the major objective is to divide a heterogeneous population into subpopulations (i.e., strata) that are homogeneous. If this is done, a relatively small sample from each stratum will provide a precise estimate of the stratum mean, and appropriate techniques for combining the stratum-mean estimates will yield a precise estimate for the population mean.

TABLE XX

## COMPARISON OF STRATIFIED AND SIMPLE RANDOM SAMPLING

Factor	Comparison
Planning of Sample	Stratified sampling requires more detailed planning and knowledge of underlying maintenance-task population than does simple random sampling.
Administration of Sampling Procedure	Stratified sampling includes all administrative aspects of simple random samples plus additional control to meet specification sample-size criteria.
Analysis of Data	Standard analytical methods are based on simple random sampling. Stratified analytical procedures for stratified samples are relatively complex and may not be available.
Sampling Efficiency	Stratified sampling generally is more efficient than simple random sampling in that variances of sample estimates are lower than for simple random samples.
Sub-hypotheses	Stratified sampling provides a means to test hypotheses on different portions of the system with adequate control. Such control is not generally available for simple random sampling.
Representativeness	Stratified sampling provides assurance that sample observations from each stratum will be observed. Simple random samples can only provide such assurance probabilistically.

In maintainability demonstration, however, known analytic techniques for stratified sampling are either inapplicable or too complex for practical use. Stratification is used in demonstration primarily to ensure representativeness; therefore, the criterion of homogeneity should be considered with this in mind. If a stratum includes a number of tasks, selection of only one or two tasks should yield a representative sample of the stratum, and across all strata the sample selections should accurately represent the total maintenance-task population.

Therefore, the tasks within a stratum should require approximately the same amount of maintenance time or the same number of man-hours, whichever is appropriate. Rigid adherence to this precept should be avoided, however. Repairing a particular electronic assembly may take approximately the same amount of time as repairing a motor generator, but the differences between the two types of actions would make it unnatural to place them in the same stratum. It seems reasonable, then, to require also that there be similarities among the tasks assigned to a stratum.

The detail to which the maintenance tasks are defined must also be considered. For example, a maintenance action may be defined as a single task, "replace Unit 01" or as two tasks, such as "replace Unit 01, which has a short" and "replace Unit 01, which has an open". Conceivably, the diagnostic times for these two tasks can differ, so that they might better be placed in different strata. There are practical limits on the detail to which tasks are defined, however -- namely; the amount of information concerning expected task times and the desire to limit the number of strata so that random-sampling analytical techniques can be reasonably applied.

There is no single correct approach to stratifying the population of maintenance-task times. The following approach is believed to be reasonable and practical:

- (1) First divide the equipment or system by physical entities, such as equipments within a system or units within an equipment. These first-level breakdowns will be called blocks.
- (2) For each block, subdivide to the highest system level at which maintenance will be performed. If the block is the highest level, no further subdivision is necessary. If an equipment is under test and the organizational-maintenance philosophy is unit replacement, subdivide to the units. In some cases, repairs or adjustments may be made within a unit, but this is considered in the next step. These elements of the subdivision will be called sub-blocks.
- (3) For each sub-block, list the associated maintenance tasks and estimated maintenance-task times or man-hours.<sup>5</sup> For a sub-block that is an LRU, removal or replacement may be the only task listed. However, if LRU adjustment or some further tasks such as crystal replacement are possible, they would also be listed as sub-block tasks.

---

<sup>5</sup>It is assumed that only tasks which relate to the specified maintainability index are listed. For example, noncritical, frequently occurring, and easily corrected malfunctions, such as light-bulb failures, would not normally be included in the task list.

- (4) Group together those tasks in each sub-block which require essentially similar actions and will be expected to have similar maintenance times or man-hours, whichever index applies. The use of historical data, the prediction effort of the maintainability-engineering group, and previous development tests should be used as inputs for the time estimates. These groups will then form part of the initial set of strata.
- (5) Compile a list of maintenance tasks that cannot be easily associated with a physical entity -- for example, repair of interconnecting cable; and tasks that are not a direct result of unreliability or degradation, such as equipment abuse and faulty installations.
- (6) Combine similar tasks of step 5 into strata, which are then added to the list of initial strata of step 4.

This scheme simply uses hardware characteristics as the first approach towards stratification and then similarity of maintenance actions and task times within hardware groups as the final criterion for stratification. The miscellaneous tasks of step 5 are added in recognition of the fact that not all maintenance actions involve the usual "black box" failures and not all actions are predictable simply from analysis of part failure rates.

This initial set of strata may have to be revised when the actual tasks to be induced and sample-size requirements are considered. Later examples illustrate this. At this point it is worthwhile to provide an illustration of this procedure.

The illustration will be concerned with a maintainability-demonstration test of an airborne doppler radar equipment consisting of the following units:

Antenna (AS)  
Receiver/Transmitter (RT)  
Frequency Tracker (FT)  
Radar Set Control (C)  
Drift Angle Indicator (ID)

The organizational-maintenance procedure for this equipment prescribes replacing all units except the receiver-transmitter, for which modularized assemblies are removed and replaced. These assemblies are as follows:

I.F. - A	Modulator (Mod)
I.F. - B	Transmitter (Tx)
Audio Amplifier (Amp)	Power Supply (PS)

Simple mechanical repairs or adjustments on the antenna and crystal replacement in the frequency trackers are also performed in the aircraft. The stratification procedure described above is shown schematically in Table XXI.

TABLE XXI

EXAMPLE OF STEP-BY-STEP STRATIFICATION

Step 1 Blocks	Step 2 Sub-blocks	Step 3 Sub-block Tasks* and Task Times	Step 4 Block Strata	Step 5 Miscellaneous Strata
Antenna	Antenna	R/R 1.0 Mech. Adjust 0.5	A - R/R A - Mech. Adjust.	
Receiver/ Transmitter	IF - A IF - B Amplifier Modulator Power Supply Transmitter	R/R 0.3 } R/R 0.3 } R/R 0.4 } R/R 0.4 } R/R 0.4 } R/R 0.5 }	IF - P/R Amp, Mod., PS - R/R Tx - R/R	
Frequency Tracker	Frequency Tracker	R/h 0.6 Replace Crystals 0.5	FT - R/R FT - Replace Crystals	
Radar Set Control	Radar Set Control	R/R 0.5	C - R/R	
Drift Angle Indicator	Drift Angle Indicator	R/R 0.5	ID - R/R	
Miscellan- eous				Repair ca- bling or con- nectors. Adjust faulty installation
*R/R - remove and replace.				

6.4.2.3 Task Occurrence Frequency and Selection

Once the initial set of strata has been established, it is necessary to estimate the frequency of occurrence of tasks in

each stratum. For tasks that result from part failure, the use of part failure rates such as those presented in MIL-HDBK-217 may be satisfactory. These failure rates, however, primarily reflect catastrophic piece-part failures and usually do not include such failure modes as degradation, part interactions, and intermittencies.

Since it is quite difficult to predict these types of failures, a simplified procedure is used whereby a failure-rate prediction based on catastrophic failures is adjusted to account for other types of malfunctions. One means for such adjustment is to analyze unit complexity.

Previous ARINC Research analysis<sup>6</sup> of historical data on electronic and electromechanical systems has indicated that the operating failure rate of a nonredundant item can be estimated by

$$\hat{\lambda} = K_e C^{1.35}$$

where

$\hat{\lambda}$  is the estimated failure rate

C is the complexity in terms of equivalent analog active elements

$K_e$  is the average failure rate of one active element operating in environment e (i.e., ground, airborne, or missile)

This type of relationship also appears in the Navy Maintainability Engineering Handbook - NAVORD OD 39223. Table XXII is abstracted from that document to provide the basis for determining C, the number of equivalent analog active elements in an equipment. The fact that the exponent of the complexity factor is greater than one is attributed in part to the interaction effects existing within the system, which can cause noncatastrophic failure occurrences such as component degradation and intermittency.

The above equation can be used in one of two ways in assessing frequency of occurrence of failure in electronic or electromechanical items:

- (1) If MIL-HDBK-217 or similar failure-rate prediction procedures are used, the total failure rate of the item can be estimated by the equation

$$\hat{\lambda} = \hat{\lambda}_{cat} C^{0.35}$$

<sup>6</sup>G. T. Bird, "On Reliability Prediction in Satellite Systems", ARINC Research Publication 4226-1-205, 1960 (see summarizing article in May 23, 1960 issue of Aviation Week).

TABLE XXII

WEIGHTING FACTORS FOR ESTIMATING EQUIVALENT ANALOG  
COMPLEXITY OF ELECTRONIC SUBSYSTEMS

AEG Type	Function	Equivalent Analog AEGs*
Transistor	Signal-level analog function	1.0
	Signal-level digital function	0.1
	Power conversion and regulation	2.0
Diodes (Semiconductor)	Signal-level analog function	0.1
	Signal-level digital function	0.01
	Power rectification	1.0
Electron Tube	Signal-level analog function	1.0
	Signal-level digital function	0.1
	Power conversion and regulation	10.0
Microwave Power Tubes	Traveling wave tubes, magnetrons, klystrons	100.0
Photoelectric Cell	Light sensor functions	0.1
Photo Multiplier	Light amplifier	10.0
Solar Cell	Power generation	0.01
Relays	General	1.0
Gyros, Position	Inertial reference	50.0**
Gyros, Rate	"Rate" signal	10.0**
Accelerometers	Acceleration measurement	1.0
Crystals	Frequency determination	1.0
Thyratron	Power switching	100.0

\* See NAVORD OD 39223, Appendix B for source of AEG data.

\*\* For short-duration missions (less than 500 hours).

where  $\hat{\lambda}_{cat}$  is the total failure-rate estimate based on catastrophic occurrences only.

- (2) If a direct prediction of total failure rate is to be obtained from complexity analysis, the following equations are used ( $\lambda$  = failures per operate hour):

Ground Systems:  $\ln \hat{\lambda}_g = -12.013 + 1.35 \ln C$

Airborne Systems:  $\ln \hat{\lambda}_a = -10.005 + 1.35 \ln C$

Missile Systems:  $\ln \hat{\lambda}_m = -6.950 + 1.35 \ln C$

For mechanical, hydraulic, and other nonelectronic items, appropriate prediction procedures presented in the literature will have to be used.<sup>7</sup> In some cases, direct failure-rate predictions are inappropriate, and the probability of task occurrence in terms of unreliability is a better measure. Such probabilities can be adjusted to an equivalent average failure rate by the relationship

$$\lambda = \frac{-\ln[1-U(t)]}{t}$$

where  $U(t)$  is the probability that the maintenance task will be required after a mission of  $t$  hours [ $U(t) = 1-R(t)$ ].

Prediction-by-function approaches are also applicable for failure-frequency estimation since they are generally based on field data that include all types of failures. Some prediction-by-function equations include complexity as a prediction parameter. The following RADC reports are applicable:

RADC-TDR-63-146 - "System Reliability Prediction by Function" (Ground Systems) - Federal Electric Corporation, May 1963, AD 406191.

<sup>7</sup>Several such sources are:

RADC-TR 68-403, "Reliability Prediction - Mechanical Stress/Strength Interference (Nonferrous)", University of Michigan, February 1969.

RADC-TR 68-114, "Data Collection for Nonelectronic Reliability Handbook", Hughes Aircraft Corporation, June 1968.

"Investigation of Reliability of Mechanical Systems", Lockheed-Georgia Company, October 1965, AD 475977.

RADC-TDR-63-300 - "System Reliability Prediction by Function" (Ground Systems, 2 Vols.), ARINC Research Corporation, August 1963, AD 481191, 481192.

RADC-TDR 65-27 "System Reliability Prediction By Function" (Ground Systems - Supplementary Report), ARINC Research Corporation, March 1965, AD 614227.

RADC-TDR 66-509 - "Avionics Reliability and Maintainability Prediction by Function," ARINC Research Corporation, October, 1966, AD 802998.

Once the failure-rate predictions are made, the relative frequency of task occurrence is calculated in a manner similar to that indicated in Table XIX of Subsection 6.4.2.1. Table XXIII shows the computations for the illustrative system.

In the table, part-failure-rate estimates corrected by the complexity factor are assumed to be used for all tasks except those involving mechanical faults. For the latter, the failure rate shown is assumed to be based on estimates of occurrence probability. Dividing the total failure rate of 830 into the individual maintenance-task strata rates yields the relative frequencies of occurrence shown in the table.

Several strata are then regrouped to yield at least a five percent frequency of occurrence. This regrouping is done to minimize the number of strata with small frequencies of occurrence, especially those which would lead to a required sample size of less than one for stratified sampling. It is not necessary to regroup if simple random sampling is to be used, in which case the relative frequencies of occurrence are used to determine task selection. If, for example, the required total sample size is 60, then 60 numbers should be drawn from a random-number table. If a random number is between 0 and 0.16, a maintenance task involving the antenna is to be simulated. If the random number is between 0.16 and 0.27, a fault resulting in the removal and replacement of an IF module is induced, etc.

For a stratified sample, the numbers in the last column of the table are the sample-size requirements. Thus, for a sample size of 60, 10 faults are to be induced involving the antenna, 7 faults are to be induced that will result in the removal and replacement of one of the IF modules, etc.

The actual faults to be induced require further analysis. For the antenna, it is seen from the relative-frequency-of-occurrence column, that removal and replacement occurs almost four

TABLE XXIII

CALCULATIONS OF RELATIVE FREQUENCY OF OCCURRENCE  
AND SAMPLE SIZE FOR RADAR DOPPLER EQUIPMENT

Maintenance Task Strata	Failure Rate 10 <sup>6</sup>	Quantity of Items	Total Failure Rate*	Relative Frequency of Occurrence, P <sub>1</sub> (Percent)	Sample Size††
Antenna R/R	105	1	105	0.127	10
Antenna-Mech/Adjust.	—	—	30**	0.036	
IF-R/R	45	2	90	0.109 = 0.11	7
Amp - R/R	10	1 }	30	0.036	3
Mod - R/R	8				
P.S. - R/R	12				
Tx - R/R	10	1	10	0.012	28
FT - R/R	400	1	400	0.482 = 0.48	
FT - Replace Crystal	20	4	80	0.096 = 0.10	6
C - R/R	35	1	35	0.042	3
ID - R/R	10	1	10	0.012	
Repair Cabling or Connector	—	—	10**	0.012	3
Adjust Faulty Installation	—	—	30**	0.036	
Total =			830	1.000	60

\*All units have 100% duty cycle. Therefore, a duty-cycle column is not shown.

\*\*Obtained from estimate of probability of occurrence, using the relationship

$$\lambda = \frac{-\ln R(T)}{T}, \quad T = \text{mission time.}$$

† Regrouped to increase all frequencies of occurrence to at least 0.05.

†† The sample sizes shown apply only to stratified sampling. For simple random sampling, the relative frequencies of occurrence are used.

times as often as the mechanical adjustment. If the maintenance task involves the antenna, the probability that it is a removal and replacement action is

$$\frac{0.127}{0.127 + 0.096} \approx 0.78$$

Therefore, a representative random sample for the antenna tasks can be obtained by selecting a random number between 0 and 1. If it is less than or equal to 0.78, induce a fault that will result in a removal and replacement. If the random number is greater than 0.78, induce a fault requiring on-board mechanical adjustment or repair.

Within these tasks, there will also be a choice of the fault to be simulated. This may involve, for example, the selection of an assembly or part and its mode of failure, and here the consideration of symptom indication may be important, especially for removal-replacement actions (see Subsection 6.4.2.4 for a sampling approach based on symptoms).

In the same fashion as discussed above with regard to the choice of antenna task, failure-mode or symptom probabilities should be analyzed to obtain their relative frequency of occurrence within a task. A random-sampling procedure is then applied in accordance with these relative probabilities to determine which faults should be induced.

For example, consider the remove-and-replace task involving the frequency tracker. From Table XXIII, 28 tasks are to be simulated for a proportional stratified sample. A failure-mode-and-effects analysis indicates that there are five major failure modes that require a remove-and-replace action and that are detectable at the unit level. These modes, their effects, and relative frequencies of occurrence are as follows:

Failure Mode	Effect	Relative Frequency of Occurrence (%)	Cumulative Range (%)
1	Inoperative	30	0 - 29.99
2	Will not lock on	20	30.00 - 49.99
3	Breaks lock	20	50.00 - 69.99
4	Drifts	15	70.00 - 84.99
5	Erratic	15	85.00 - 100.00

To determine which faults to induce for simulating the 28 remove-and-replace tasks, 28 four-digit numbers are selected from a random-number table. If the number selected is between 0 and 2999, then a fault or malfunction that makes the frequency tracker inoperative must be induced. A random number between 3000 and 4999 would indicate that a fault resulting in inability to lock-on effect is to be simulated, etc.

Within any failure mode, a choice would then have to be made concerning the specific means of fault inducement (e.g., which lead to disconnect or which part to replace with a known faulty part). If, with respect to the maintenance action, there is no discernible difference, the simplest means can be used. If, however, the fault selection can affect maintenance time (e.g., disconnecting one lead may cause secondary symptoms, while disconnecting another will not), then, again, a random-selection criterion is advisable.

To minimize the biasing problems due to task rehearsals and the occasional problem of not being able physically to induce the required fault, it is prudent to select a much larger number of possible tasks than required. For example, if the required sample size is 60, a total of 120 tasks may be selected initially and procedures for fault inducement established. When the test is actually run, half of these prepared tasks are then selected for actual observation.

#### 6.4.2.4 Stratification and Task Selection Based on the Symptom Matrix

Another approach to developing strata and selecting tasks is to use the symptom-matrix formulation presented in Volume I of this study. The symptom matrix of concern here is one that relates unit malfunctions, associated symptoms, and occurrence probabilities.

If symptoms are considered rather than maintenance tasks only, recognition is given to the fact that for many equipments, especially those with easily replaceable units, a large portion of the corrective maintenance action involves fault location; therefore, the use of symptom information rather than task information might be a better basis for stratification.

The stratification can be accomplished in one of three basic ways:

- (1) Stratify by symptoms, and sample unit malfunctions within symptoms
- (2) Stratify by units, and sample symptoms within unit malfunctions

(3) Stratify by symptom/unit malfunction combinations

To illustrate these three approaches, the symptom matrix developed for the APN-147 Doppler Radar Set monitored in this study is shown in Table XXIV.<sup>8</sup> The entries in the cell representing the  $i^{\text{th}}$  unit and the  $j^{\text{th}}$  symptom are described as follows:

$i^{\text{th}}$ Unit	$j^{\text{th}}$ symptom	
	$\lambda_{1j}$	$P(U_1 S_j)$
	$P(S_j U_1)$	$P(U_1S_j)$

where

$\lambda_{1j}$  = the failure rate of elements in the  $i^{\text{th}}$  unit whose malfunction will result in the  $j^{\text{th}}$  symptom

$P(S_j|U_1)$  = the probability that the  $j^{\text{th}}$  symptom will occur given the  $i^{\text{th}}$  unit malfunctions

$P(U_1|S_j)$  = the probability that the  $i^{\text{th}}$  unit has malfunctioned given the  $j^{\text{th}}$  symptom appears

$P(U_1S_j)$  = the probability of the joint occurrence of an  $i^{\text{th}}$  unit malfunction and  $j^{\text{th}}$  symptom

One difficulty in using this approach is to develop the  $\lambda_{1j}$  estimates since this requires fairly detailed analysis of the system design. Volume I presents a discussion of this estimating process.

The following relationships exist under the assumption that concurrent failures do not occur:

$$P(S_j|U_1) = \frac{\lambda_{1j}}{\sum_k \lambda_{1k}} = \frac{\lambda_{1j}}{\lambda_{U_1}}$$

<sup>8</sup>Only major units are shown. Such items as mounting trays and cabling, which can and do sometimes cause failure, are not included in the example.

TABLE XXIV  
SYNPTOM MATRIX FOR APN147 DOPPLER RAJAR

		Symptoms		Presented Symptoms						
		1	2	4	7	1 and 4	1, 4 and 7	2 and 5	3 and 6	$P(u_1)$
Replacable Unit										
Receiver-Transmitter						127.000 1.000 0.293 0.223				0.22
Frequency Tracker	28 0.631 0.050	3 2 0.008 0.149 0.005	27 0.074 0.784 0.047	9 5 0.026 0.871 0.016	272.000 0.750 0.628 0.478	13 0.035 1.0 0.022	9 3 0.025 1.0 0.016			0.63
Antenna	9 2 0.166 0.503 0.016	18 0.333 0.851 0.032			28 0.500 0.063 0.348					0.10
Indicator	1 4 0.333 0.031 0.002		1 4 0.333 0.042 0.002	1 4 0.333 0.229 0.002						0.01
Radar Set Control	6 0 0.301 0.133 0.010		6 0 0.200 0.174 0.010		6 000 0.300 0.013 0.010	2 0 0.100 1 0 0.003				0.04
$P(S_j)$	0.08	0.04	0.06	0.02	0.76	0.02	0	0.02	0.02	1.00

$$\begin{matrix} \text{Cell Entry} & 1 & j \\ \lambda_{1j} & P(S_j|u_1) \\ P(u_1|S_j) & P(u_2|S_j) \end{matrix}$$

$$P(U_1|S_j) = \frac{\lambda_{1j}}{\sum_h \lambda_{hj}} = \frac{\lambda_{1j}}{\lambda_{S_j}}$$

$$P(U_1 S_j) = \frac{\lambda_{1j}}{\sum_{hk} \lambda_{hk}} = \frac{\lambda_{1j}}{\lambda_T}$$

$$P(S_j) = \frac{\sum_h \lambda_{hj}}{\sum_{hk} \lambda_{hk}} = \frac{\lambda_{S_j}}{\lambda_T}$$

$$P(U_1) = \frac{\sum_k \lambda_{1k}}{\sum_{hk} \lambda_{hk}} = \frac{\lambda_{U_1}}{\lambda_T}$$

where

$\lambda_{U_1}$  = the failure rate of the 1<sup>th</sup> unit

$\lambda_{S_j}$  = the failure rate of all items whose malfunction will produce the j<sup>th</sup> symptom

$\lambda_T$  = the total equipment failure rate

For simplicity, assume that  $n = 100$  maintenance actions are to be induced by stratified sampling. Under the three methods, the sampling described below would take place.<sup>9</sup>

- (1) Symptom Stratification. The number of tasks in which the j<sup>th</sup> symptom is induced is  $P(S_j) \cdot n$ . Hence the task sample sizes by symptom are as follows:

<u>Symptom</u>	<u>Sample Size</u>	<u>Symptom</u>	<u>Sample Size</u>
1	8	1 and 4	76
2	4	1, 4, and 7	0
4	6	2 and 5	2
7	2	3 and 6	2

<sup>9</sup>For simple random sampling, the procedure is analagous to that discussed for maintenance-task stratification.

Within each symptom, malfunctions are introduced randomly into the units according to the probabilities  $P(U_i|S_j)$ . Thus for generating symptom  $S_1$ , there should be a 0.63 probability of inducing the symptom-generating fault in the frequency tracker, a 0.20 probability for inducing it in the antenna, a 0.03 probability for inducing it in the indicator, and a 0.13 probability for inducing it in the control unit.

- (2) Unit Stratification. The number of tasks in which the  $i^{\text{th}}$  unit is induced is  $P(U_i) \cdot n$ . Hence, the task sample sizes by unit-failure inducements are as follows:

<u>Unit</u>	<u>Sample Size</u>
RT	22
FT	63
AS	10
ID	1
C	4

Within each unit, malfunctions are introduced randomly according to the probabilities  $P(S_j|U_i)$ . Thus for the antenna there should be a 0.17 probability of inducing a fault that will generate  $S_1$ , a 0.33 probability for  $S_2$ , and a 0.50 probability for the  $S_1$  and  $S_4$  symptom combination.

- (3) Symptom/Unit Malfunction Stratification. For this case, each unique symptom/unit malfunction combination represents a stratum. All the strata are thus combinations with non-zero entries for  $P(U_i S_j)$ , and the sample sizes are  $P(U_i S_j) \cdot n$ . Thus 22 samples are to be selected with symptoms 1 and 4 with an RT malfunction, 5 samples are to be selected with symptom 1 with a malfunction in the frequency tracker, etc.

The choice between the three basic approaches described above depends on the expected comparative representativeness and on sampling difficulties. The third method is probably least practical since it will generally entail too fine a breakdown, requiring regrouping of some combinations. Method 2 is probably preferable to Method 1 because Method 1 may place too much emphasis on the symptom aspects of maintenance.

Method 2 is actually the same initial approach as the task-sampling procedure described earlier. The difference is that for this method the actual faults to be introduced within a unit are governed by symptom occurrence rather than failure occurrence.

The symptom approach can be recommended over the task-sampling approach in the following circumstances:

- Fault location and, to a lesser extent, checkout time are expected to account for the most significant portion of the corrective maintenance action.
- The necessary information and resources (personnel, time, money) for developing the symptom matrix are available.

## SECTION VII

### STATISTICAL MAINTAINABILITY-DEMONSTRATION PLANS

#### 7.1 GENERAL

This section reviews the maintainability-demonstration test methods of MIL-STD-471 and then presents a number of other tests applicable to maintainability demonstration. Each of the plans 1 through 4 of MIL-STD-471, which are the most frequently applied procedures, has one or more analogs in the set of alternative plans presented herein.

Four different categories of alternatives are presented:

- (1) Fixed-sample-size tests, lognormal distribution
- (2) Sequential tests, lognormal distribution
- (3) Nonparametric tests
- (4) Bayesian tests

For convenience, the following standardized format is used to describe the alternative plans for the non-Bayesian tests:

- . General Description of Test
- . Underlying Assumptions
- . Hypothesis
- . Sample Size
- . Decision Procedure
- . Discussion

Guidelines for selecting among the alternatives are also presented.

#### 7.2 MIL-STD-471 PLANS

##### 7.2.1 General

The six test methods of MIL-STD-471 are reviewed in this subsection. These test methods include situations covering various types of maintainability specifications, test procedures, and underlying assumptions. Table XXV summarizes the various forms of maintainability parameters that are tested by five of the six MIL-STD test methods under consideration. Test method 5 is not actually a decision test but an approach to estimating the percentage of maintenance tasks between the observed sample extremes; therefore, it is not listed in the table, but it is reviewed.

TABLE XXV

## SUMMARY OF MIL-STD-471 SPECIFIED PARAMETERS

Test Method	$\bar{M}_{ct}$	$\bar{M}_{pt}$	$\bar{M}$	$M_{\max ct}$	$M_{\max pt}$	$\tilde{M}_{ct}$	$\tilde{M}_{pt}$
1	X			X			
2	X	X	X	X			
3						X	
4				X	X	X	X
6		X			X		

$\bar{M}_{ct}$  = Mean corrective-maintenance downtime

$\bar{M}_{pt}$  = Mean preventive-maintenance downtime

$\bar{M}$  = Mean maintenance downtime consisting of corrective and preventive in the same time period

$M_{\max ct}$  =  $p^{th}$  percentile of corrective-maintenance downtime

$M_{\max pt}$  =  $p^{th}$  percentile of preventive-maintenance downtime

$\tilde{M}_{ct}$  = Median corrective-maintenance downtime

$\tilde{M}_{pt}$  = Median preventive-maintenance downtime

The six methods are reviewed below with respect to type of test, conditions of use, sample sizes, and accept/reject criteria. The reader is advised to consult the standard for the detailed test procedures.

#### 7 2 Test Method 1

Type of Test. Two sequential tests are performed -- one for  $\bar{M}_{ct}$  and one for  $M_{\max ct}$ . Both tests must be passed for an accept decision. The producer's and consumer's risks are held to a maximum of 16 percent.

Conditions of Use.  $\bar{M}_{ct}$  and  $M_{\max ct}$  must both be specified. A lognormal distribution of corrective-maintenance times is

assumed, although it is claimed that test risks are only slightly changed if the actual distribution is exponential or normal. The specified  $\bar{M}_{ct}$  must be greater than 10 minutes and less than 100 minutes, and the ratio  $M_{max\ ct}/\bar{M}_{ct}$  must be less than 3.  $M_{max\ ct}$  can be either the 90th or 95th percentile.

Sample Size. The sequential tests are truncated at 100 maintenance actions. Fewer than that number will generally be required before a decision is reached.

Accept/Reject Criteria. The number of maintenance times less than and greater than the specified  $\bar{M}_{ct}$  and  $M_{max\ ct}$  are recorded and compared with tabulated values of the accept and reject numbers for the two sequential tests. A decision rule for cases in which no decision has been made after 100 observed actions is also provided.

Discussion. The Standard calls for stratified sampling, which may seriously affect the risks of the sequential plan unless the order of the sampling is strictly random. It is emphasized that the specified maintainability requirements ( $\bar{M}_{ct}$  and  $M_{max\ ct}$ ) are the unacceptable levels -- that is, equipment which exactly meets these levels will have a low probability of passing the test. It is also emphasized that the procedure is based on converting the mean and percentile specifications to equivalent specifications of binomial parameters, leading to the use of sequential tests when the binomial distribution applies.

### 7.2.3 Test Method 2

Type of Test. Fixed-sample test employing the central-limit theorem for sample statistics of  $\bar{M}_{ct}$ ,  $\bar{M}_{pt}$ ,  $\bar{M}$ , and  $M_{max\ ct}$ . Only consumer's risks are considered; the value of  $\beta$  is contractually determined.

Conditions of Use. To demonstrate  $M_{max\ ct}$ , a lognormal repair-time distribution must be assumed. When both preventive- and corrective-maintenance indexes are specified, the proportion of each type of maintenance during a representative operational period must be estimated.

Sample Size. A minimum of 50 corrective-maintenance tasks and 50 preventive-maintenance tasks (if  $\bar{M}_{pt}$  is specified) must be observed.

Accept/Reject Criteria. Equations are given for estimating  $\bar{M}_{ct}$ ,  $\bar{M}_{pt}$ , and  $M_{max\ ct}$ . Accept/reject values for each of these

indices are then computed from the formulas presented. They are a function of the sample sizes, the specified  $\beta$  risks, and, for  $M_{\max ct}$ , the value of the  $p$ th percentile of interest. If the estimated  $\tilde{M}_{ct}$ ,  $\tilde{M}_{pt}$ ,  $\tilde{M}$ , and  $M_{\max ct}$  are less than the corresponding critical values, the equipment passes the test.

Discussion. Notice 1 to MIL-STD-471, dated 9 April 1968, corrected an erroneous statement that the risk probability was related to the producer's risk. The test for  $M_{\max ct}$  simply compares the estimated  $M_{\max ct}$  to that specified; therefore, the risk for this portion of the test is not as stated in the standard but is approximately 50 percent.

#### 7.2.4 Test Method 3

Type of Test. Fixed-sample test, assuming a lognormal distribution of repair times.

Conditions of Use. The specified value of  $\tilde{M}_{ct}$ , say  $\tilde{M}_0$ , the median corrective-maintenance time, is calculated from an equation given in the standard based on an assumed standard deviation of 0.55 and a known value of  $\tilde{M}_{ct}$ , say  $\tilde{M}_1$ , for which only a 5-percent acceptance probability ( $\beta$  risk) is desired. The accept/reject criterion is based on a producer's risk of 5 percent.

Sample Size. A sample size of 20 maintenance actions is required.

Accept/Reject Criteria. Equations are presented for obtaining the statistics necessary to test for acceptable median repair times. The standard "t" test equation (using logarithms of repair time) is given to ensure that there is a 95-percent chance of accepting equipments with  $\tilde{M}_{ct} = \tilde{M}_0$  and only a 5-percent chance of accepting equipments with  $\tilde{M}_{ct} = \tilde{M}_1$ .

Discussion. The description and notation of the test method in the standard can be somewhat confusing. Actually, the basic requirement that is assumed to exist is the unacceptable value of the median, which is denoted by  $ERT_{\max}$ . This value corresponds to a  $\beta$  risk of 5 percent (the Standard erroneously states 10 percent). The sample size of 20 is fixed and, assuming  $J$  (log-base 10) = 0.55, the value of the median that corresponds to a 5-percent producer's risk (denoted by  $ERT_{\text{spec}}$  in the standard) is given by the equation  $ERT_{\text{spec}} = 0.37 ERT_{\max}$ . It is emphasized that  $ERT_{\text{spec}}$  is completely determined by the sample size of 20,

the assumed  $\sigma$ , and the known  $ERT_{max}$ , and is not based on prior consideration of system or logistic requirements. Therefore, it is misleading to ascribe the word "specified" to  $ERT_{spec}$ . The better way to evaluate test applicability is the  $ERT_{max}$  specification and corresponding  $\beta$  risk.

#### 7.2.5 Test Method 4

Type of Test. Nonparametric test for proportion to demonstrate achievement of specified  $\bar{M}_{ct}$  or  $\bar{M}_{pt}$  and  $M_{max\ ct}$  or  $M_{max\ pt}$ .

Conditions of Use. No underlying distribution of maintenance times is assumed. Both a median and maximum time must be specified. Test criteria are given separately for the median ( $\bar{M}$ ) and for the 95th percentile ( $M_{max}$ ). The individual median and  $M_{max}$  tests are made at either the 75-percent or 90-percent confidence level.

Sample Size. The sample size is 50 each for corrective or preventive maintenance tasks.

Accept/Reject Criteria. For both the median and  $M_{max}$  tests, critical values are provided for the number of maintenance actions greater than specified values for a sample size of 50 and for the 75-percent and 90-percent confidence levels. An accept decision is made only if both the median and  $M_{max}$  tests are passed.

Discussion. The individual median and  $M_{max}$  tests are equivalent to tests on a binomial parameter. It is important to note that the specified parameters represent unacceptable maintainability levels; e.g., the 90-percent-confidence-level test of  $\bar{M}_{ct}$  is such that if the median corrective-maintenance time is equal to that specified, there is only a 10-percent chance of passing the test. Thus this test is based on specified  $\beta$  risks (1-confidence level) corresponding to the specified maintainability level for median or 95th percentile. The Beta risk corresponding to an equipment accept decision (both individual tests are passed) is approximately 1.5 percent for the 90-percent-confidence-level tests and 9 percent for the 75-percent-confidence-level tests.

#### 7.2.6 Test Method 5

Type of Test. A nonparametric procedure for obtaining a confidence-interval estimate of the proportion of maintenance-task times in the population that will be included between the observed sample extremes. Tolerance interval is the usual terminology for such an estimate.

Conditions of Use. No underlying distribution of maintenance times is assumed.

Sample Size. MIL-STD-471 presents the required sample size for three confidence levels (90, 95, and 99) and three population percentages (90, 95, and 99). For example, the standard shows that a sample size of 47 is required to be 95-percent confident that 90 percent of the population's maintenance-task times will be contained within the observed sample extremes.

Accept/Reject Criteria. No such criteria exist for this test. The test results provide a measure of spread, and this measure may be used to determine acceptability.

Discussion. Since this procedure is basically one of estimation and since it is not geared to a particular demonstration requirement, it should be used more for informative purposes than for decision purposes.

#### 7.2.7 Test Method 6

Type of Test. Nonstatistical-type test. Sample statistics are compared with the specified values.

Conditions of Use. Values of  $\bar{M}_{pt}$  or  $M_{max pt}$ , or both, must be specified. The percentile point defining the latter may be any value. The frequency of occurrence for each type of preventive-maintenance task must be estimated.

Sample Size. Not specified. All preventive-maintenance tasks are to be performed.

Accept/Reject Criteria. An average value for  $\bar{M}_{pt}$  is estimated from the data on the basis of actual maintenance times weighted by frequency of occurrence. This value is then compared with the specified  $\bar{M}_{pt}$  to determine acceptance or rejection. To test for  $M_{max pt}$ , the observed preventive-maintenance times are ordered. The observed  $p^{th}$  percentile time is compared with the specified  $M_{max pt}$  to determine conformance.

Discussion. This method provides no risk control for either producer or consumer and is therefore limited in its application. Such a procedure is better suited to obtaining information than to decision-making.

## 7.3 ALTERNATIVE NON-BAYESIAN TESTS

### 7.3.1 General

This subsection is devoted to the presentation of various alternative tests which, to varying degrees, parallel the MIL-STD-471 plans but offer greater flexibility in risk assignment, test parameters, hypotheses, and form of testing. Discussion of Bayesian tests is deferred until Subsection 7.5. The tests presented are standard statistical tests or adaptations thereof. However, little or no experience with their application to maintainability demonstration has been accumulated. They are believed to be practical for this purpose, but final judgment must await more extensive use.

### 7.3.2 Comments on Sampling Procedure and Sample Size

#### 7.3.2.1 Sampling Procedure

A general comment concerning the type of sampling is in order. All tests (both MIL-STD-471 and the alternative plans) employ analytical procedures that are based on the assumption of simple random sampling. If proportional stratified sampling is used, large sample sizes and a relatively small number of strata are required to ensure validity. In most cases, however, the assumption of simple random sampling is conservative in the sense that the sample size based on this assumption is larger than would be required if analytical procedures were based on stratified sampling errors. For sequential tests, it is recommended that only simple random sampling be performed.

#### 7.3.2.2 Sample Size

Tests for which only a single risk ( $\alpha$  or  $\beta$ ) is specified do not generally have associated sample-size requirements. Since the risk of a wrong decision generally decreases with increasing sample size, the larger the value of the sample size, the greater the assurance of a correct decision.

As a very general rule, it would be appropriate to specify a minimum sample size on the order of 25 to 30 even for cases in which a sample-size equation indicates that a lower value would be acceptable. For very complex systems involving hundreds of different types of maintenance actions, it will be desirable to specify a larger sample size to provide reasonable assurance of representativeness.

### 7.3.3 Lognormal Distribution

Since the lognormal distribution has often been an adequate representation of various repair-type distributions, many of the test procedures are based on a lognormal-distribution assumption

concerning the maintenance characteristic of interest. The reader is advised to review the characteristics of the lognormal distribution presented in Appendix A.

#### 7.3.4 Notation

The following notation is used:

$X$  = the maintenance-characteristic random variable

$\mu = E(X)$ , the mean value of  $X$

$X_p$  = the  $(1-p)^{th}$  percentile of  $X$

$\tilde{M} = X_{0.50}$ , the median of  $X$

$d^2 = E(X-\mu)^2$ , the variance of  $X$

$\theta = E(\ln X)$ , the mean value of  $\ln X$

$\sigma^2 = E(\ln X - \theta)^2$ , the variance of  $\ln X$

$Z_p$  = the standardized normal deviate exceeded with proba-

bility  $p$ , i.e.,  $\int_{Z_p}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = p$

For the lognormal distribution, the density is

$$f(X) = \frac{1}{\sqrt{2\pi} \sigma X} e^{-\frac{1}{2\sigma^2} (\ln X - \theta)^2}, \quad 0 < X < \infty$$

where  $\theta = E[\ln X]$  and  $\sigma^2 = \text{variance} [\ln(X)]$

Then for the lognormal distribution,

$$\mu = e^{\theta + \sigma^2/2}$$

$$\tilde{M} = e^{\theta}$$

$$X_p = e^{\theta + Z_p \sigma}$$

$$d^2 = e^{2\theta + \sigma^2} (e^{\sigma^2} - 1)$$

#### 7.3.5 Prior Estimation of $\sigma^2$

Several of the tests (1, 2, 3, 7, and the Bayesian test) require that an estimate of  $\sigma^2$  be used for determining specified values, evaluating the sample size, or developing the decision criterion. To aid in obtaining such estimates, the corrective-maintenance data collected for this study were analyzed.

Table XXVI presents the observed  $\sigma^2$  values for the natural logarithm of active-corrective-maintenance time ( $\sigma^2_{\ln ART}$ ) and maintenance man-hours per active corrective maintenance action ( $\sigma^2_{\ln MH}$ ). Time spent on administrative or logistic activities is not included, since these time elements would not normally be considered in a maintainability demonstration. Variances are shown for two cases: no-trouble-found actions included, and no-trouble-found actions excluded. The latter is generally more pertinent for demonstration in which faults are to be induced, while the former may be more applicable for natural-failure tests.

The sample sizes from which the variances were calculated are shown so that a weighted average of several variance values can be used if it is determined that such an average would be best for estimating the variance of the equipment under consideration.<sup>10</sup>

In addition, multiple-linear-regression analysis was employed to develop prediction equations of the form

$$\hat{\sigma}^2 = b_0 + b_1X_1 + b_2X_2 + \dots + b_rX_r$$

where

$\hat{\sigma}^2$  is the predicted value of  $\sigma^2$  (the dependent variable)

$X_i$  is the  $i^{\text{th}}$  prediction parameter (independent variable)

$b_0, b_1, \dots, b_r$  are computed regression coefficients

The least-squares method was applied through a computerized procedure to obtain the regression coefficients, as well as other regression and correlation measures such as the standard error of estimate, simple, partial, and multiple correlation coefficients, and  $t$  and  $F$  statistics for significance tests.

Significance tests were used to determine the statistical significance of each parameter at the nominal 20-percent significance level. For efficiency, a step-wise procedure was used.

<sup>10</sup>In MIL-STD-471, Test Method 3, a value of  $\sigma_{\log_{10} x} = 0.55$  is re-

commended for use. On converting to base-10 logarithms, the average standard deviation of the 21 equipments in Table XXVI for no-trouble-found actions included is approximately 0.38. Since the value in the Standard was developed from data obtained a number of years ago, if the two estimates are representative of the same populations, a desirable trend towards reduced variation is evident. It can be conjectured that this variance reduction is due in part to greater emphasis on maintainability, leading to such features as automatic test equipment, computerized troubleshooting, and modular or unit replacement.

TABLE XXVI  
OBSERVED VARIANCES — AIR FORCE AVIONIC  
AND GROUND EQUIPMENT

Equipment	No-Trouble-Found Actions Included			No-Trouble-Found Actions Excluded		
	Number of Observations	$\sigma^2_{lnART}$	$\sigma^2_{lnMH}$	Number of Observations	$\sigma^2_{lnART}$	$\sigma^2_{lnMH}$
<b>Avionic Equipment</b>						
APQ-110-Terrain Radar	25	0.584	0.691	16	0.629	0.681
APN-167-Padar Altimeter	5	0.588	1.388	<5	--	--
AJQ-20-Inertial Bomb Nav.	42	0.373	0.590	22	0.233	0.365
APQ-113-Fine Control Radar	16	0.814	0.808	8	1.593	1.496
ARC-109 UHF Transceiver	11	0.430	0.562	10	0.417	0.546
APS-109 Radar Homing and Warning	5	0.810	1.02	<5	--	--
APN 59 Search Radar	40	1.100	1.680	35	1.034	1.598
APN-147 Doppler Radar	6	0.956	1.484	<5	--	--
ASN-35 Doppler Computer	9	0.283	0.528	6	0.197	0.378
ARN-21 TACAN	33	0.739	0.986	27	0.567	0.732
APN-157 LORAN	18	0.165	0.237	14	0.208	0.228
ASN-24 Navigation Computer	14	0.885	1.147	10	0.861	1.053
ADF-73 Automatic Direction Finder	9	0.699	0.574	6	0.594	0.594
ARC-90 UHF Communications	5	0.995	1.197	<5	--	--
<b>Ground Equipment</b>						
GYK-4 Data Processor	12	1.378	2.091	10	1.073	1.54
GSA-51 Display Consoles	24	1.188	1.638	24	1.188	1.64
GSH-12 Recorder/Reproducer	11	1.259	1.555	11	1.259	1.55
GSQ-72 Punch Card System	7	0.521	0.412	7	0.521	0.412
FPS-27 Transmitter	43	0.996	1.426	42	0.984	1.376
FPS-27 Receiver	24	0.656	1.184	23	0.645	1.094
FPS-27 ECM	9	0.505	0.602	8	0.346	0.601

This procedure<sup>11</sup> selects the most significant of the independent variables, as measured by the simple-correlation coefficients, and then determines if the addition to the overall correlation of the most likely remaining candidate (determined by analyzing semipartial correlation coefficients) is significant. If there is a significant candidate, it is included, and the process is repeated until none of the remaining variables can add significantly to the multiple-correlation coefficient.

Because the dependent variable  $\sigma^2$  is an average value based on samples of individual repair times, it is reasonable to weight each observed value of  $\sigma^2$  in proportion to the sample size. Rigorous weighting would involve consideration of the variance of a variance estimate to achieve required homoscedasticity for significance tests (assuming that necessary distributional assumptions are satisfied).

For practical purposes, however, weighting was done simply by the square root of the number of sample observations for each equipment, since this tends to prevent disproportionate weighting and past experience using this approach with reliability- and maintainability-prediction-by-function analysis has proven favorable.

More than 20 possible prediction parameters were evaluated. Included were various information-theory parameters, design and complexity parameters, and maintenance design and concept parameters. It was decided to eliminate from consideration any calculated regression equation that did not have logical justification. This primarily involved checking the sign of the computed regression coefficient. For example, it would seem logical that a parameter highly positively correlated with complexity should have a positive coefficient, indicating that as the parameter value increases, so does the variance of repair time.

The presence of "wrong" signs does not invalidate an equation, because such a result may be due to complex interactions among the selected parameters. However, past experience has shown that predictions based on such equations are considered by some as being of doubtful value. Since the results obtained in this analysis with "good" signs were favorable, the question became somewhat academic.

Three equations were developed:

- A. Equations for the variance of the natural logarithm of active corrective-maintenance time -- no-trouble-found actions excluded (see note in Table XXVIII for including such actions).

<sup>11</sup>The ARINC Research regression program uses the Square Root Method described in "A Square Root Method of Selecting a Minimum Set of Variables in Multiple Regression", *Psychometrika*, Vol. 16, No. 3, A. Summerfield and A. Lubin, September 1951.

- (1) Prediction inputs -- basic design and maintenance parameters. Sample size = 21 equipments (290 maintenance actions).
- (2) Prediction inputs -- basic design and maintenance parameters plus symptom-matrix parameters. Sample size = 13 equipments (176 maintenance actions).

B. Equation for the variance of the natural logarithms of active corrective maintenance manhours -- no-trouble-found actions excluded (see note in Table XXVIII for including such actions).

Prediction inputs -- basic design and maintenance parameters. Sample size = 21 equipments (290 maintenance actions).

The difference between A(1) and A(2) is that the latter requires development of a symptom matrix from which the efficiency of information transmission, ET, can be obtained. Information was available to compute ET values for only 13 of the 21 systems. Equation A(1) does not include symptom-matrix parameters.

For application to corrective-maintenance time, equation A(2) will generally provide more precise results, but quantification of the ET parameter is somewhat laborious. If a symptom matrix has not been developed for predicting fault-location time as prescribed by the procedure presented in Volume 1, A(1) may be preferred in terms of the effort involved.

The prediction parameters involved in the three equations are summarized and quantified in Table XXVII. Table XXVIII presents the three prediction equations and necessary parameter constraints to avoid invalid application. The list of sample equipments presented in Table XXVI should also be examined to determine the applicability of the sample equipments to the equipments under consideration. The actual regression runs are duplicated in Appendix B.

Various regression and correlation statistics pertaining to the three equations are summarized in Table XXIX.  $R^2$ , the multiple correlation coefficient squared, is a measure of the total variation in the observed variance values that can be explained by the variation in the prediction parameters. Since the calculated F values exceed the corresponding critical F values, these  $R^2$  values are statistically significant at the 0.05 level. The standard error of estimate,  $s$ , is the standard deviation of the observed (sample) values from the regression plane. As described below,  $s$  is used for confidence-interval prediction.

TABLE XXVII  
PREDICTION-PARAMETER QUANTIFICATION

Symbol	Description	Equation	Quantification
ET	<u>Efficiency of Information Transmission</u> - A measure of the amount of information presented by failure symptoms	A(2)	See Volume 1, Section 2
MC	<u>Maintenance Complexity Factor</u> - A checklist score representing the relative amount of effort and time required for the preparation and fault-location activities	A(2)	See note below
RP	<u>Relative Power Consumption</u> - The steady-state power, in watts, consumed by the equipment in its most power-consuming mode of operation	A(1) A(2) B	1 - Low (0 to 250 watts)  2 - Medium (250 watts to 5 kilowatts)  3 - High (over 5 kilowatts)
SD	<u>Signal-Data Handling</u> - The type of circuitry used to process a signal	A(1) B	Use sum of following applicable weights:  Analog (Signal) - 1 Analog (Computing) - 2 Digital - 2
TC	<u>Test Concept</u> - The equipment-testing philosophy for on-line maintenance	A(1) B	Automatic Self Test - 1 Semi-Automatic Self Test - 2 External Test Set - 3 Standard Test Provisions - 4  (Use average if more than one category applies)

NOTE: To obtain the value of MC, applicable test characteristics from the following list are checked, and the checked scores or weights (shown in parentheses) are totaled.

**Test-Equipment Requirements**

- None (1)
- One equipment (4)
- Two equipments (5)
- Three equipments (6)

**Maintenance-Manual Requirements**

- None (1)
- One manual (4)
- Two or more manuals (5)

**Tool Requirements**

- None (1)
- Standard tool kit (2)
- Special tools (3)
- Both standard and special tools (4)

**Special-Handling-Equipment Requirements**

- None (1)
- Equipment required (3)

**Access for Preparation**

- No need for preparatory action (1)
- Removal of panel or plate (3)

**Symptom Indications**

- Obvious symptoms present (1)
- Some indications present (4)
- No clear failure indications (6)

**Warm-Up Requirements**

- None (1)
- Two to five minutes' warm-up (3)
- More than five minutes' warm-up (6)

**Cycle-Down Requirements**

- None (1)
- Two to five minutes' cycle-down (3)
- More than five minutes' cycle-down (6)

TABLE XXVIII

## VARIANCE PREDICTION EQUATIONS\*

Equation A(1): $\hat{\sigma}_{\ln ART}^2 = -0.800 + 0.385(RP) + 0.221(SD) + 0.117(TC)$	
<u>Constraints:</u> $4 \leq RP + SD + TC \leq 10$	
Equation A(2): $\hat{\sigma}_{\ln ART}^2 = 0.036 + 0.225(RP) + 0.023(MC) - 0.785(ET)$	
<u>Constraints:</u> $12 \leq MC \leq 30$ $0.07 \leq ET \leq 0.70$	
Equation B: $\hat{\sigma}_{\ln MH}^2 = -0.711 + 0.501(RP) + 0.170(SD) + 0.133(TC)$	
<u>Constraints:</u> $4 \leq RP + SD + TC \leq 10$	
<p>*Equations apply for predicting the variance when no-trouble-found actions are excluded. To predict <math>\sigma^2</math> when no-trouble-found (NTF) actions are included, the following average relationship derived from the data in Table XXVI can be used:</p> $\hat{\sigma}_{NTF-IN}^2 = (1.106)\hat{\sigma}_{NTF-OUT}^2$	

TABLE XXIX

## SUMMARY OF REGRESSION STATISTICS

Equation	Dependent Variable	Sample Size	R	R <sup>2</sup>	Standard Error of Estimate, s	Calculated F Value	Critical F Value (5%)
A(1)	$\sigma_{\ln ART}^2$	21	0.81	0.65	0.256	10.65	3.20
A(2)	$\sigma_{\ln ART}^2$	13	0.85	0.72	0.210	7.82	3.86
B	$\sigma_{\ln MH}^2$	21	0.80	0.64	0.301	10.15	3.20

To obtain a confidence-interval prediction, the equations for the confidence limits are as follows:

$$\begin{aligned}\sigma^2_L &= \hat{\sigma}^2 - k \\ \sigma^2_U &= \hat{\sigma}^2 + k\end{aligned}$$

where

$\hat{\sigma}^2$  is the predicted value of  $\sigma^2$  obtained from the regression equation

$\sigma^2_L$  and  $\sigma^2_U$  are lower and upper confidence limits, respectively, for the variance

$$k = t_{(\alpha/2, d)} s \left( \frac{1}{m} + \sum_{i=1}^r c_{i1} x_1^2 + 2 \sum_{j=1+1}^r \sum_{i=1}^{r-1} c_{ij} x_1 x_j \right)^{1/2}$$

where

$m$  = number of observations used in the regression analysis

$r$  = number of independent variables (prediction parameters in the equation)

$t_{(\alpha/2, d)}$  =  $t$  statistic for a 100  $(1-\alpha)\%$  two-sided confidence interval based on  $d = (m-r-1)$  degrees of freedom (use  $t_{(\alpha, d)}$  for a one-sided interval)

$s$  = standard error of estimate

$c_{ij}$  = computed Gauss multiplier

$x_i$  = deviation  $(X_i - \bar{X}_i)$  where  $X_i$  is the value for the  $i^{\text{th}}$  independent variable in the prediction equation and  $\bar{X}_i$  is the mean of the observed values for the  $i^{\text{th}}$  parameter in the sample data used for the regression analysis

Table XXX presents the information necessary for computing  $k$ .

Prediction of confidence limits might be desirable for investigative purposes. For example, a two-sided interval provides a range for  $\sigma^2$  that would provide information on minimum and maximum expected sample sizes for a particular demonstration test under

TABLE XXX  
FACTORS FOR CONFIDENCE-INTERVAL PREDICTION

Equation	Dependent Variable	m	r	$t_{(.05,d)}$ (90% Limits)	s	Independent Variable		$X_1$	Gauss Multipliers
						Symbol	1		
A(1)	$\sigma^2_{\ln ART}$	21	3	1.74	0.2559	RP	1	2.071	$c_{11} = 0.1011$
						SD	2	2.130	$c_{12} = 0.00927$
						TC	3	2.457	$c_{13} = 0.01501$
									$c_{22} = 0.05615$
									$c_{23} = 0.01845$
									$c_{33} = 0.1021$
A(2)	$\sigma^2_{\ln ART}$	13	3	1.83	0.2097	RP	1	1.951	$c_{11} = 0.3256$
						MC	2	1.898	$c_{12} = -0.01531$
						ET	3	0.402	$c_{13} = 0.2043$
									$c_{22} = 0.00256$
									$c_{23} = -0.01105$
									$c_{33} = 0.3231$
B	$\sigma^2_{\ln MH}$	21	3	1.74	0.3015	RP	1	2.071	$c_{11} = 0.1011$
						SD	2	2.130	$c_{12} = 0.00927$
						TC	3	2.457	$c_{13} = 0.01501$
									$c_{22} = 0.05615$
									$c_{23} = 0.01845$
									$c_{33} = 0.1021$

consideration. An upper confidence limit may be computed on  $\sigma^2$  and used in the demonstration as a conservative practice. In Section 7.5, specific uses of confidence limits on  $\sigma^2$  are discussed in connection with developing a prior distribution of  $\theta$ , the expected value of  $\ln X$ .

### 7.3.6 Fixed-Sample-Size Tests, Lognormal Distribution

#### 7.3.6.1 Introduction

The tests presented in this section are those in which the sample size is specified and the decision criterion is based on a lognormal assumption for the distribution of  $X$ , the maintenance-time random variable of interest. In the interest of generality,  $X$  will not normally be further identified, but it should be clear that  $X$  can represent either corrective- or preventive-maintenance time or man-hours, and that parameters such as mean, median, or percentile refer to the specific time identification represented by  $X$ .

#### 7.3.6.2 Test Number 1: Test on Mean of Lognormal Distribution

General Description of Test. A fixed-sample test on the mean of a lognormal distribution that is based on the asymptotic normality of the sample arithmetic mean (central-limit theorem). Sample sizes of 20 or more should be chosen to approximate the asymptotic distribution adequately.

Underlying Assumptions. Maintenance time can be adequately described by a lognormal distribution. The sample size is large enough (say, at least 20) so that the central-limit theorem provides a good normal approximation to the distribution of the sample mean.

Hypotheses.  $H_0$ : Mean =  $\mu_0$

$H_1$ : Mean =  $\mu_1$ , ( $\mu_1 > \mu_0$ )

Sample Size. For a test with producer's risk  $\alpha$  and consumer's risk  $\beta$ ,

$$n = \frac{(Z_{\alpha\mu_0} + Z_{\beta\mu_1})^2}{(\mu_1 - \mu_0)^2} (e^{\hat{\sigma}^2} - 1)$$

where  $\hat{\sigma}^2$  is a prior estimate of the variance of the logarithms of maintenance times.

$Z_{\alpha}$  and  $Z_{\beta}$  are standardized normal deviates.

Decision Procedure. Obtain a random sample of  $n$  maintenance times,  $X_1, X_2, \dots, X_n$ , and compute the sample arithmetic mean,

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

and the sample variance

$$\hat{d}^2 = \frac{1}{n-1} \left( \sum_{i=1}^n X_i^2 - n\bar{X}^2 \right)$$

Accept  $H_0$  if  $\bar{X} \leq \mu_0 + Z_\alpha \hat{d}/\sqrt{n}$

Reject  $H_0$  otherwise.

Discussion. This test corresponds to the mean test of Method 2, MIL-STD-471, except that it provides a control on both  $\alpha$  and  $\beta$  through proper choice of the sample-size value.

Since the mean and variance of  $X$  are assumed to be finite, by the central-limit theorem, the sample arithmetic mean,  $\bar{X}$ , is approximately normal for large  $n$  with mean  $= E(X)$ , and variance  $= V(X)$ . Hence, on the basis that

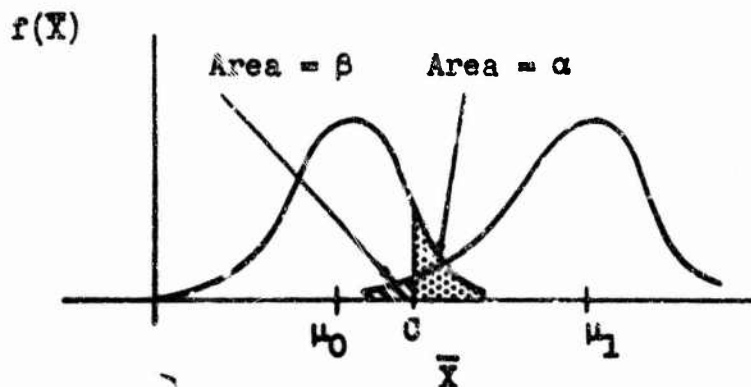
$$V(X) = d^2 = e^{2\theta + \sigma^2} (e^{\sigma^2} - 1) = \mu^2 (e^{\sigma^2} - 1),$$

the following are true:

$$\text{under } H_0: \bar{X} \sim N(\mu_0, \mu_0^2 (e^{\sigma^2} - 1)/n)$$

$$\text{under } H_1: \bar{X} \sim N(\mu_1, \mu_1^2 (e^{\sigma^2} - 1)/n)$$

The distribution of  $\bar{X}$  under the two hypotheses can be represented by the following diagram:



It is necessary to choose a critical value  $C$  ( $H_0$  rejected if  $\bar{X} > C$ ) such that if  $\mu = \mu_0$ , there is a probability of  $\alpha$  that  $\bar{X} > C$ , and if  $\mu = \mu_1$ , the probability that  $\bar{X} \leq C = \beta$

Thus, under the asymptotic normality of  $\bar{X}$ ,  $n$  can be determined from the following two equations:

$$\mu_0 + z_\alpha \sqrt{v(\bar{X}_n)} = C$$

$$\mu_1 - z_\beta \sqrt{v(\bar{X}_n)} = C$$

or

$$\mu_0 + z_\alpha \mu_0 (e^{\sigma^2} - 1)^{1/2} / \sqrt{n} = \mu_1 - z_\beta \mu_1 (e^{\sigma^2} - 1)^{1/2} / \sqrt{n}$$

which yields

$$n = \frac{(z_\alpha \mu_0 + z_\beta \mu_1)^2}{(\mu_1 - \mu_0)^2} (e^{\sigma^2} - 1)$$

To provide maximum assurance that the  $\alpha$  and  $\beta$  risks are being preserved when a prior estimate  $\hat{\sigma}^2$  is used in the above equation, the upper limit of any interval prediction or estimate on  $\sigma^2$  should be used.

To indicate the rapid approach to normality of the sample mean  $\bar{X}$ , 100 samples each of sizes 10, 20, 30, and 40 were generated from a lognormal distribution with parameters  $\theta = 3.75$  and  $\sigma = 0.85$ . These values are approximately the average values observed in the field data obtained for 21 different Air Force equipments. For each sample of size  $n$ ,  $\bar{X}$  was computed and the distribution of 100 such values was tested by the Kolmogorov-Smirnov Test. Table XXXI compares the theoretical and observed values of the mean and standard deviation of the  $\bar{X}_n$  distributions and the Kolmogorov-Smirnov statistic,  $D$ , of the maximum absolute deviation between the observed distribution function and the theoretical normal distribution function.

For 100 samples of  $\bar{X}_n$ , the hypothesis of normality is rejected at the 10-percent significance level if  $D > 0.122$ ; at the 5-percent significance level, the critical region is  $D > 0.136$ . From Table XXXI none of the samples can be rejected at the 0.05 level, but the  $n = 20$  case would have been rejected for  $\alpha = 0.10$ . Surprisingly, the  $n = 10$  case yielded the smallest value of  $D$ , but this must

be interpreted as being due to sampling irregularities since it is known that the approach to normality of any  $\bar{X}$  distribution is monotonic in the sense that as  $n$  increases, the observed distribution more closely approaches the theoretical normal.

TABLE XXXI

THEORETICAL AND OBSERVED VALUES OF MEANS AND VARIANCES  
AND KOLMOGOROV-SMIRNOV STATISTICS FOR NORMALITY TEST

Sample Size, $n$	Theoretical Mean	Observed Mean	Theoretical $\sigma^2 \bar{X}$	Observed $\sigma^2 \bar{X}$	Observed Maximum Deviation, $D$
10	61.0	62.0	19.9	19.0	0.072
20	61.0	60.3	14.0	12.9	0.131
30	61.0	61.8	11.5	11.1	0.098
40	61.0	60.2	9.93	10.0	0.119

Example. It is desired to test the hypothesis that the mean corrective-maintenance time is equal to 30 minutes against the alternative hypothesis that  $M_{ct} = 45$  minutes under the lognormal assumption. Previous data indicate that  $\sigma^2$  (the variance of  $\ln X$ )  $\approx 0.6$ . The  $\alpha$  and  $\beta$  risks are set at 5 percent.

Then:  $\mu_0 = 30$

$\mu_1 = 45$

If the sample-size equation is used, the following is obtained:

$$n = \frac{[1.645(30) + 1.645(45)]^2}{(45-30)^2} (e^{0.6}-1) = 56$$

### 7.3.6.3 Test Number 2: Test on Median of Lognormal Distribution

General Description of Test. An exact fixed-sample test on the median of a lognormal distribution that is based on the  $t$  distribution. Since the test is exact, sample size is restricted only by the  $\alpha$  and  $\beta$  risks. Since the median of a lognormal is equal to  $e^{\theta}$ , tests on  $\bar{M}$  are equivalent to tests on  $\theta$ .

Underlying Assumptions. Maintenance times can be adequately described by a lognormal distribution.

Hypotheses.  $H_0$ : Median =  $\tilde{M}_0$  or  $\theta = \theta_0 = \ln \tilde{M}_0$

$H_1$ : Median =  $\tilde{M}_1$  or  $\theta = \theta_1 = \ln \tilde{M}_1$

Sample Size. The sample-size equation presented below for given  $\alpha$  and  $\beta$  risks is derived in Reference 1 (see list of references at end of the Section).

$$n = 1 + \left( \frac{1}{a} + \frac{Z_\alpha^2}{2} \right) \quad (\text{Round up to next integer.})$$

where

$$a = \frac{(\theta_1 - \theta_0)^2}{\tilde{\sigma}^2 (Z_\alpha + Z_\beta)^2}, \quad \tilde{\sigma}^2 = \text{prior estimate of } \sigma^2$$

Decision Procedures. Compute

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n \ln X_i \quad \text{and} \quad s = \frac{1}{n-1} \left[ \sum_{i=1}^n (\ln X_i)^2 - n\hat{\theta}^2 \right]$$

Accept  $H_0$  if  $\hat{\theta} \leq \theta_0 + t_{n-1, \alpha} s/\sqrt{n}$ ; reject  $H_0$  otherwise, where  $t_{n-1, \alpha}$  is the 100  $(1-\alpha)$ th percentile of the  $t$  distribution with  $(n-1)$  degrees of freedom.

Discussion. This test corresponds to Test 3 of MIL-STD-471 except that control on both  $\alpha$  and  $\beta$  is provided through proper choice of  $n$ . Since  $\tilde{M} = e^\theta$ , any hypothesis on  $\tilde{M}$  yields an equivalent hypothesis on  $\theta$  by the relationship  $\theta = \ln \tilde{M}$ . Since  $\theta = E(\ln X)$  and  $\ln X$  is normally distributed with mean  $\theta$  and variance  $\sigma^2$ , the quantity

$$\frac{\hat{\theta} - \theta}{s/\sqrt{n}}$$

has the  $t$  distribution with  $(n-1)$  degrees of freedom.

Thus lognormality and a median specification yield the well known  $t$  test based on a normal distribution. The sample-size equations are based on the approximate normality of  $s$ . Details of the nature of the approximation are given in Reference 1.

#### 7.3.6.4 Test Number 3: Test on Critical Percentile

General Description of Test. A fixed-sample-size test on the maintenance time corresponding to a particular percentile when the underlying distribution is lognormal. The decision criterion is based on the asymptotic normality of the maximum-likelihood estimate of a percentile value.

Underlying Assumptions. Maintenance times can be adequately described by a lognormal distribution. Under this assumption, it is further assumed that sample size is large enough (say greater than 10) so that the statistic  $(\hat{\theta} + Z_p s)$  is approximately normally distributed.

Hypotheses.  $H_0$ :  $(1-p)^{th}$  percentile,  $X_p = T_0$  or

$$P [X > T_0] = p$$

$H_1$ :  $(1-p)^{th}$  percentile,  $X_p = T_1$  or

$$P [X > T_1] = p, (T_1 > T_0)$$

Sample Size. To meet specified  $\alpha$  and  $\beta$  risks, the sample size to be used is given by the formula

$$n = \left( \frac{2 + Z_p^2}{2} \right) \hat{\sigma}^2 \left( \frac{Z_\alpha + Z_\beta}{T_1 - T_0} \right)^2 \quad (\text{Round up to next integer.})$$

where

$\hat{\sigma}^2$  is a prior estimate of  $\sigma^2$

$Z_p$  is the normal deviate corresponding to the  $(1 - p)^{th}$  percentile

Decision Procedure. Compute

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n \ln X_i$$

$$X^* = T_0 + Z_\alpha s \left[ \frac{1}{n} + \frac{Z_p^2}{2(n-1)} \right]^{1/2}$$

where

$$s^2 = \frac{1}{n-1} \left[ \sum_{i=1}^n (\ln X_i)^2 - n\hat{\theta}^2 \right]$$

Accept  $H_0$  if  $\hat{\theta} + Z_p s \leq X^*$ .

Reject  $H_0$  otherwise.

Discussion. This test corresponds to the test of  $M_{\max}$  of Test 2, MIL-STD-471. As pointed out in Subsection 7.2.3, the MIL-STD-471 test has an unspecified risk and its use is thus quite limited.

The basis for this test is as follows: Under the lognormal assumption, the  $(1-p)^{\text{th}}$  percentile value is given by  $X_p = e^{\theta + Z_p \sigma}$ .

Taking logarithms gives  $\ln X_p = \theta + Z_p \sigma$ ; therefore, if maximum-likelihood estimates are used for the normal parameters  $\theta$  and  $\sigma$ , the maximum-likelihood estimate of the  $(1-p)^{\text{th}}$  percentile is

$$\ln \hat{X}_p = \hat{\theta} + Z_p s$$

where

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n \ln X_i$$

$$s^2 = \frac{1}{n-1} \left[ \sum_{i=1}^n (\ln X_i)^2 - n\hat{\theta}^2 \right]$$

As shown in reference 2, p. 58,  $\hat{X}_p$  is approximately normal unless  $n$  is quite small (say, under 5), with mean and variance given as follows:

$$E(\hat{X}_p) = E[\hat{\theta} + Z_p s] = \theta + Z_p \sigma = X_p$$

$$\text{Var}(\hat{X}_p) = \text{Var}[\hat{\theta} + Z_p s] = \sigma^2 \left[ \frac{1}{n} + \frac{Z_p^2}{2(n-1)} \right]$$

To meet the producer's and consumer's risk requirements, a critical value  $X^*$  has to be chosen for the sample estimate of the  $(1-p)^{\text{th}}$  percentile,  $X_p$ , such that

$$P [\hat{X}_p > X^* | X_p = T_0] = \alpha \text{ and}$$

$$P [\hat{X}_p > X^* | X_p = T_1] = 1 - \beta$$

Because of the approximate normality of  $\hat{X}_p$ ,  $X^*$  must satisfy the equation

$$\frac{X^* - T_0}{\sqrt{\text{Var } \hat{X}_p}} = Z_\alpha \text{ or } X^* = T_0 + Z_\alpha \sqrt{\text{Var } \hat{X}_p}$$

and, similarly,

$$\frac{X^* - T_1}{\sqrt{\text{Var } \hat{X}_p}} = Z_{1-\beta} \text{ or } X^* = T_1 - Z_\beta \sqrt{\text{Var } \hat{X}_p}$$

Equating the right-hand sides of the above two equations and substituting for  $\text{Var } \hat{X}_p$  (using  $n$  rather than  $n-1$  in the second term) yields the sample-size equation

$$n = \left( \frac{2 + Z_p^2}{2} \right) \sigma^2 \left( \frac{Z_\alpha + Z_\beta}{T_1 - T_0} \right)^2$$

**Example.** Assume that the following two hypotheses are to be tested with  $\alpha = \beta = 0.10$ :

$$H_0: X_{0.05} = 1.5$$

$$H_1: X_{0.05} = 2.0$$

Assume that  $\sigma^2$  is estimated to be 1.

Then

$$\begin{aligned}
 n &= \left( \frac{2 + Z_p^2}{2} \right) \sigma^2 \left( \frac{Z_\alpha + Z_\beta}{T_1 - T_0} \right)^2 \\
 &= \left[ 1 + \frac{(1.645)^2}{2} \right] \left( \frac{2.56}{0.5} \right)^2 \\
 &= 62
 \end{aligned}$$

and the critical value is

$$\begin{aligned}
 X^* &= T_0 + Z_\alpha s \left[ \frac{1}{n} + \frac{Z_p^2}{2(n-1)} \right]^{1/2} \\
 &= 1.5 + 1.28s \left[ \frac{1}{62} + \frac{(1.645)^2}{122} \right]^{1/2} \\
 &= 1.5 + 0.250s
 \end{aligned}$$

#### 7.3.6.5 Test Number 4: Test on Critical Maintenance Time

General Description of Test. A fixed-sample-size test on the probability that a maintenance action will be completed before a specified time interval. This type of test is applicable when direct control on availability or turnaround time is important. For example, if operational requirements dictate that an aircraft be available within 30 minutes after an initial alert, the probability of completing any necessary equipment repairs within that time period should be high.

Underlying Assumptions. Maintenance times can be adequately described by a lognormal distribution. With the lognormality assumption, it is further assumed that the statistic  $(\hat{\theta} + Z_p s)$  is approximately normally distributed.

Hypotheses.  $H_0: P[X > T] = p_0$  or  $T = X_{p_0}$

$H_1: P[X > T] = p_1$  or  $T = X_{p_1}, (p_1 > p_0)$

where  $T$  is a specified time and

$X_{p_1}$  is the  $(1-p_1)^{th}$  percentile,  $i = 0$  or  $1$

Note that this test and the preceding test have the same null hypothesis  $H_0$  but differ in  $H_1$  since test number 3 keeps the percentile value,  $X_p$ , constant and varies the time,  $T$ , while this test keeps  $T$  constant and varies  $X_p$ .

Sample Size. To meet specified  $\alpha$  and  $\beta$  risk requirements, the sample size is determined from the equation

$$n = \left( \frac{k^2 + 2}{2} \right) \left( \frac{Z_\alpha + Z_\beta}{Z_{p_0} - Z_{p_1}} \right)^2 \quad (\text{Round up to next integer.})$$

where

$$k = \frac{Z_\alpha Z_{p_1} + Z_\beta Z_{p_0}}{Z_\alpha + Z_\beta}$$

and the Z's represent standardized normal deviates.

Decision Procedure. Compute

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n \ln X_i$$

$$s^2 = \frac{1}{n} \left[ \sum_{i=1}^n (\ln X_i)^2 - n\hat{\theta}^2 \right]$$

$$k = \frac{Z_\alpha Z_{p_1} + Z_\beta Z_{p_0}}{Z_\alpha + Z_\beta}$$

Accept  $H_0$  if  $\hat{\theta} + ks \leq T$ .

Reject  $H_0$  otherwise.

Discussion. This test is an application of the well known plans for acceptance sampling by variables where the quality criterion is based on percent defective. A discussion of this test and derivation of the equations for  $k$  and  $n$  are given in Reference 3, pp. 303-311.

Example. It is desired that 95 percent of all repairs be completed within 1.5 hours. An equipment for which only 85 percent of repairs are completed in less than 1.5 hours is considered unacceptable. A test is to be conducted with  $\alpha = \beta = 0.10$ .

The hypotheses are:

$$H_0: X_{0.05} = 1.5$$

$$H_1: X_{0.15} = 1.5$$

Then

$$k = \frac{Z_{0.10}Z_{0.15} + Z_{0.10}Z_{0.05}}{Z_{0.10} + Z_{0.10}} = \frac{(1.28)(1.04) + (1.28)(1.645)}{2.56} = 1.34$$

and

$$n = \left( \frac{(1.34)^2 + 2}{2} \right) \left( \frac{2.56}{1.645 - 1.04} \right)^2$$

$$= 34$$

Thus 34 observations are sampled and the requirement is demonstrated if

$$\hat{\theta} + ks = \hat{\theta} + 1.34s \leq 1.5$$

#### 7.3.6.6 Test Number 5: Test on Joint Specifications of Lognormal Parameters

General Description of Test. A fixed-sample-size test of any one of various possible pairs of lognormal parameters such as the mean and percentile. Table XXXII presents the equations for converting the values of the two parameters specified to an equivalent specification on  $\theta$  and  $\sigma^2$ , the mean and variance of  $\ln X$ , a normally distributed random variable. Since the form of the critical region depends on the relationship between the null and alternative parameter values, the reasonable restriction that  $\sigma_1^2 > \sigma_0^2$  was made to simplify the presentation.

Assumptions. The distribution of maintenance times can be described adequately by a lognormal distribution.

Hypotheses. The null and alternative hypotheses may specify values for any of the following pairs of parameters:

Mean - Variance; Mean - Percentile; Median - Percentile;

$P_1^{\text{th}}$  Percentile -  $P_2^{\text{th}}$  Percentile; Mean - Median

Table XXXII presents the equations for converting hypothesized values of any of the above pairs to an equivalent hypothesis on  $\theta$  and  $\sigma^2$ . The general set of hypotheses is then as follows:

$$H_0: \theta = \theta_0, \sigma^2 = \sigma_0^2$$

$$H_1: \theta = \theta_1, \sigma^2 = \sigma_1^2, (\sigma_1^2 > \sigma_0^2)$$

TABLE XXXII

RELATIONSHIP BETWEEN  $\theta$  AND  $\sigma^2$  TO PAIRS OF  
SPECIFIED LOGNORMAL PARAMETERS

Specified Lognormal Parameters	Equivalent Specification on Normal Parameters
Mean = $\mu$  Variance = $d^2$	$\theta = \ln \mu^2 - \frac{1}{2} \ln(d^2 + \mu^2)$  $\sigma^2 = \ln \left[ \frac{\mu^2 + d^2}{\mu^2} \right]$
Mean = $\mu$  $p^{\text{th}}$ Percentile = $X_p$ (see Note)	$\theta = \ln \mu - \sigma^2/2$  $\sigma^2 = \left[ Z_p - \left( Z_p^2 + 2 \ln(\mu/X_p) \right)^{1/2} \right]^2$
Median = $\tilde{M}$  $p^{\text{th}}$ Percentile = $X_p$	$\theta = \ln \tilde{M}$  $\sigma^2 = \left[ \frac{\ln(X_p/\tilde{M})}{Z_p} \right]^2$
$(1-p_1)^{\text{th}}$ Percentile = $X_{p_1}$  $(1-p_2)^{\text{th}}$ Percentile = $X_{p_2}$	$\theta = \frac{Z_{p_2} \ln X_{p_1} - Z_{p_1} \ln X_{p_2}}{Z_{p_2} - Z_{p_1}}$  $\sigma^2 = \left( \frac{\ln X_{p_2} - \ln X_{p_1}}{Z_{p_2} - Z_{p_1}} \right)^2$
Mean = $\mu$  Median = $\tilde{M}$	$\theta = \ln \tilde{M}$  $\sigma^2 = 2 \ln(\mu/\tilde{M})$
NOTE: The Mean-Percentile specification does not lead to a unique lognormal distribution. Equations for $\theta$ and $\sigma^2$ represent the more reasonable of two possible parameter sets. For application, the following inequality must be satisfied: $Z_p^2 > 2 \ln(X_p/\mu)$ .	

with associated risks of  $\alpha$  and  $\beta$ . For the common mean-percentile test it can be determined from Table XXXII that the requirement  $\sigma_1^2 > \sigma_0^2$  means that  $X_{p_1}/\mu_1 > X_{p_0}/\mu_0$ .

**Sample Size.** The sample size for this test to meet specified  $\alpha$  and  $\beta$  risks involves the distribution of the noncentral chi-square distribution, for which tables are not generally available. It can be shown for this application that if the sample size is 25 or more, a reasonable normal approximation to the noncentral chi-square distribution is possible. Since sample sizes of 25 or greater have been established as standard for all tests described in this report, this restriction is not critical.

The equations for obtaining  $n$  are as follows:

$$K = (\theta_1 - \theta_0)^2 / \sigma_0^2, \quad V = \sigma_1^2 / \sigma_0^2$$

$$M_0 = K / (V - 1)^2, \quad M_1 = VK / (V - 1)^2$$

$$A_0 = (1 + 2M_0) / (1 + M_0), \quad A_1 = V(1 + 2M_1) / (1 + M_1)$$

$$B_0 = 2(1 + M_0), \quad B_1 = 2V(1 + M_1)$$

$$F = B_0 + B_1, \quad G = A_0 + A_1 + (A_0^{1/2} Z_\alpha - A_1^{1/2} Z_{1-\beta})^2$$

$$a = (B_0 - B_1)^2, \quad b = 4(A_1 B_0 + A_0 B_1) - 2FG, \quad c = G^2 - 4A_0 A_1$$

Then the required sample size is given by the equation

$$n = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

Table XXXIII presents the solution of these equations for various values of  $K$  and  $V$  for the cases  $\alpha = \beta = 0.05, 0.10$ , and  $0.20$

To use the table, the  $K$  and  $V$  values are calculated from the hypothesized values of  $\theta_0, \theta_1, \sigma_0^2$ , and  $\sigma_1^2$ . For the appropriate risk (0.05, 0.10, or 0.20), the value of  $n$  can be determined from the table by interpolation. To be conservative, for nonlisted  $K$  and  $V$  values, the next higher  $V$  and next lower  $K$  values in the table may be used to ensure that the risks are no greater than specified.

TABLE XXVIII

SAMPLE SIZES FOR TEST NUMBER 5

FOR:  $\alpha = \beta = 0.05, 0.10, 0.20$ , AND FOR VARIOUS  $K$  AND  $V$  VALUES

$\alpha = 0.05, \beta = 0.05$																					
$K \backslash V$	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0	$K \backslash V$	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0.02	456	311	209	149	111	87	71	59	51	44	0.12	91	85	78	69	61	54	48	43	39	35
0.04	326	246	175	133	103	82	68	57	49	43	0.14	79	75	70	63	56	50	45	41	37	34
0.04	253	204	157	121	96	78	65	55	48	42	0.16	69	67	63	57	52	47	42	39	35	32
0.05	207	174	139	111	89	74	62	53	46	41	0.18	62	60	57	53	48	44	40	37	34	31
0.06	175	152	125	102	84	70	60	52	45	40	0.20	56	55	52	49	45	41	38	35	32	30
0.07	152	133	114	95	79	67	57	50	44	39	0.25	45	45	43	41	39	36	34	32	30	28
0.08	134	121	104	88	75	64	55	48	43	38	0.30	38	38	37	36	34	32	30	29	27	25
0.09	120	110	96	83	71	61	53	47	42	37	0.35	32	32	32	31	30	29	28	26	25	25
0.10	109	97	89	78	67	59	51	45	41	37	0.40	28	28	28	28	27	26	25	25	25	25
$\alpha = 0.10, \beta = 0.10$																					
$K \backslash V$	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0	$K \backslash V$	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0.01	467	256	153	102	74	57	46	38	32	28	0.12	56	53	48	42	38	33	30	26	25	25
0.02	278	189	128	91	68	53	44	36	31	27	0.14	48	46	42	38	34	31	28	25	25	25
0.03	198	150	109	81	63	50	42	35	30	27	0.16	42	41	38	35	32	29	26	25	25	25
0.04	154	124	96	74	55	48	40	34	29	26	0.18	38	37	35	32	30	27	25	25	25	25
0.05	126	106	85	68	55	45	38	33	29	25	0.20	34	33	32	30	28	25	25	25	25	25
0.07	87	78	66	55	47	40	35	30	27	25	0.22	31	31	30	28	26	25	25	25	25	25
0.10	66	61	55	48	41	36	32	28	25	25	0.24	29	28	27	26	25	25	25	25	25	25
$\alpha = 0.20, \beta = 0.20$																					
$K \backslash V$	1.05	1.10	1.15	1.20	1.25	1.30	1.35	1.40	1.45	1.50	$K \backslash V$	1.05	1.10	1.15	1.20	1.25	1.30	1.35	1.40	1.45	1.50
0.005	466	304	197	134	97	74	58	48	40	34	0.040	71	67	61	54	47	42	37	32	29	25
0.010	258	201	149	111	85	67	54	45	38	33	0.045	63	60	55	50	44	39	35	31	28	25
0.015	179	150	120	94	75	61	50	42	36	31	0.050	57	55	51	46	41	37	33	30	27	25
0.020	137	120	100	82	67	55	47	40	34	30	0.060	48	46	44	40	37	33	30	27	25	25
0.025	111	100	86	72	61	51	44	37	33	29	0.070	41	40	38	36	33	30	28	25	25	25
0.030	93	86	76	65	56	48	41	36	31	28	0.080	36	35	34	32	30	28	26	25	25	25
0.035	80	75	67	59	51	44	39	34	30	27	0.090	32	32	31	29	27	26	25	25	25	25
											0.100	29	29	28	27	25	25	25	25	25	25

Note -- A line under 25 (i.e.,  $\underline{25}$ ) signifies that a lower sample size than 25 may be used but 25 is recommended as a minimum for  $M$  demonstration.

Decision Procedure. The decision procedure is as follows:

(1) Obtain a sample of  $n$  observations of maintenance times,

$$X_1, X_2, \dots, X_n.$$

$$(2) \text{ Compute } Z = \frac{1}{\sigma_0^2} \sum_{i=1}^n (\ln X_i - \rho)^2 = \frac{n}{\sigma_0^2} [s^2 + (\bar{Y} - \rho)^2]$$

$$\text{where } \rho = \frac{\theta_0 \sigma_1^2 - \theta_1 \sigma_0^2}{\sigma_1^2 - \sigma_0^2} = \frac{e_0 V - \theta_1}{V - 1}$$

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n \ln X_i$$

$$s^2 = \frac{1}{n} \left[ \sum_{i=1}^n (\ln X_i)^2 - n \bar{Y}^2 \right]$$

(3) Compute

$$M_0 = \frac{(\theta_0 - \rho)^2}{\sigma_0^2} = \frac{K}{(V - 1)^2},$$

$$u = \frac{1 + 2M_0}{1 + M_0}, \quad v = n \frac{(1 + M_0)^2}{1 + 2M_0}$$

(4) Compute

$$C = \frac{u}{2} (Z_\alpha + \sqrt{2v - 1})^2$$

(5) If  $Z \geq C$ ,  $H_0$  is rejected

If  $Z < C$ ,  $H_0$  is accepted

Discussion.<sup>12</sup> The basis of this test is developed from the Neyman-Pearson lemma, which states that the best critical region for rejecting  $H_0$  consists of points in the sample space such that

$$\frac{L(\underline{X}|H_0)}{L(\underline{X}|H_1)} \leq k_\alpha$$

<sup>12</sup> For further details see M. Kendall and H. Stuart, Advanced Theory of Statistics, Vol. 2, Griffin, 1961, pp. 174-176, 227-229.

where

$\underline{X} = (X_1, X_2, \dots, X_n)$ , the sample observations

$L(\underline{X}|H_j)$  is the likelihood of the  $n$  sample observations if  $H_j (j=0,1)$  is true

$k_\alpha$  is a constant ( $>0$ ) such that the  $\alpha$ -risk is satisfied

If  $Y_1 = \ln X_1$ , then

$$L(\underline{Y}|H_j) = (\sqrt{2\pi} \sigma_j)^{-n} \exp\left[-\frac{1}{2\sigma_j^2} \sum_{i=1}^n (Y_i - \theta_j)^2\right], j=0,1$$

and for  $\sigma_1^2 > \sigma_0^2$ , the above likelihood-ratio inequality can be shown to reduce to

$$\sum_{i=1}^n (Y_i - \rho)^2 \geq c_\alpha$$

where  $c_\alpha$  is a constant independent of the observations.

Since  $Y \sim N(\theta_0, \sigma_0^2)$  under  $H_0$ ,  $(Y - \rho)/\sigma_0 \sim N[(\theta_0 - \rho)/\sigma_0, 1]$

Then it can be shown that  $\sum_{i=1}^n \left(\frac{Y_i - \rho}{\sigma_0}\right)^2$  has a noncentral chi-

square distribution with  $n$  degrees of freedom and noncentrality parameter  $n \left(\frac{\theta_0 - \rho}{\sigma_0}\right)^2 = nM_0$ . Thus a test of size  $\alpha$  is obtained

by finding the  $100 \times (1 - \alpha)$  percentile of a noncentral chi-square distribution with  $n$  degrees of freedom and noncentrality parameter  $nM_0$ . If this percentile point is denoted by  $\chi_n'^2(nM_0, \alpha)$ ,

the decision criterion is as follows:

$$\text{Reject } H_0 \text{ if } \sum_{i=1}^n \left(\frac{Y_i - \rho}{\sigma_0}\right)^2 = Z \geq \chi_n'^2(nM_0, \alpha)$$

where

$$P[Z \geq \chi_n'^2(nM_0, \alpha) | H_0] = \alpha$$

and

$Z \sim \chi_n'^2(nM_0)$ , a noncentral chi-square variate.

The power of the test is then defined as

$$1 - \beta = P\left[Z \geq \chi_n'^2(nM_0, \alpha) | H_1\right]$$

and thus the distribution of  $Z$  under  $H_1$  is required. Since  $(Y - \rho) \sim N(\theta_1 - \rho, \sigma_1^2)$  under  $H_1$ , then

$$\sum_{i=1}^n \left( \frac{Y_i - \rho}{\sigma_1} \right)^2 = \frac{\sigma_0^2}{\sigma_1^2} Z \sim \chi_n'^2(nM_1), \text{ where } M_1 = \left( \frac{\theta_1 - \rho}{\sigma_1} \right)^2$$

Hence, the power function is

$$\begin{aligned} 1 - \beta &= P\left[ \frac{\sigma_0^2}{\sigma_1^2} Z \geq \frac{\sigma_0^2}{\sigma_1^2} \chi_n'^2(nM_0, \alpha) | H_1 \right] \\ &= P\left[ \chi_n'^2(nM_1) \geq \frac{1}{V} \chi_n'^2(nM_0, \alpha) | H_1 \right] \end{aligned}$$

If  $\chi_n'^2(\lambda)$  is a noncentral chi-square variate with  $n$  degrees of freedom and noncentrality parameter  $\lambda$ , then let

$$u = \frac{n + 2\lambda}{n + \lambda}, \quad v = \frac{(n + \lambda)^2}{n + 2\lambda}$$

Then for  $v$  over 30,

$$\left( \frac{2\chi_n'^2(\lambda)}{u} \right)^{1/2} - (2v - 1)^{1/2} \sim N(0, 1) \text{ approximately.}$$

From this approximation, the following are obtained from the  $\alpha$  and  $1 - \beta$  equations

$$\chi_n'^2(nM_0, \alpha) = \frac{u_0}{2} (Z_\alpha + \sqrt{2v_0 - 1})^2$$

$$\chi_n'^2(nM_0, \alpha) = v \frac{u_1}{2} (Z_{1-\beta} + \sqrt{2v_1 - 1})^2$$

where

$$u_1 = \frac{1 + 2M_1}{1 + M_1}, \quad v_1 = n \frac{(1 + M_1)^2}{1 + 2M_1}, \quad i = 0, 1$$

Equating the right-hand sides of the above two equations leads to the equation for  $n$ .

Example. Assume that

$H_0$ : Mean = 0.5, 95th Percentile = 1.5 hrs.

$H_1$ : Mean = 0.8, 95th Percentile = 2.5 hrs.

$$\alpha = \beta = 0.05$$

Equivalently,

$$H_0: \theta = -1.128, \sigma^2 = 0.869$$

$$H_1: \theta = -0.715, \sigma^2 = 0.984$$

Then

$$K = \frac{(\theta_1 - \theta_0)^2}{\sigma_0^2} = \frac{(-0.715 + 1.128)^2}{0.869} \approx 0.20$$

$$V = \frac{\sigma_1^2}{\sigma_0^2} = \frac{0.984}{0.869} \approx 1.13$$

From Table XXXIII, for  $\alpha = \beta = 0.05$ ,  $K = 0.20$ , and  $V = 1.1$ , the required sample size is found to be equal to 56 observations.

### 7.3.7 Sequential Tests - Lognormal Distribution

#### 7.3.7.1 Introduction

Three different sequential maintainability-demonstration tests are presented in this subsection:

- Test on a joint specification of lognormal parameters
- Test on the mean, median, or percentile of a lognormal distribution
- Test on a critical maintenance time,  $\sigma^2$  unknown

### 7.3.7.2 Test Number 6: Sequential Test on Joint Specification of Lognormal Parameters

**Description of Test.** This test is the sequential analog of Test Number 5, in which any one of various possible pairs of lognormal parameters may be specified. Table XXXII is used again to convert the original specifications to one in terms of  $\theta$  and  $\sigma^2$ , the mean and variance of  $\ln X$ , which is normally distributed.

**Assumptions.** Maintenance times can be adequately described by a lognormal distribution. Simple random sampling is performed. The original specification is such that  $\sigma_1^2 > \sigma_0^2$ .

**Hypotheses.** Any of the following pairs of parameters may be specified:

Mean - Variance; Mean - Percentile; Median - Percentile

$P_1^{\text{th}}$  Percentile -  $P_2^{\text{th}}$  Percentile; Mean - Median

Table XXXII presents the equations for converting hypothesized values of any of the above pairs to an equivalent hypothesis on  $\theta$  and  $\sigma^2$ . The general set of hypotheses is then as follows:

$$H_0: \theta = \theta_0, \sigma^2 = \sigma_0^2$$

$$H_1: \theta = \theta_1, \sigma^2 = \sigma_1^2 \quad (\sigma_1^2 > \sigma_0^2)$$

with specified  $\alpha$  and  $\beta$  risks.

**Sample Size.** The sample size for a sequential test is a random variable. For the hypothesized pairs, the expected values of  $n$  are as follows:

$$E(n|\theta = \theta_0, \sigma^2 = \sigma_0^2) = \frac{2[(1 - \alpha) \ln B + \alpha \ln A]}{\ln\left(\frac{\sigma_0^2}{\sigma_1^2}\right) + \frac{1}{\sigma_1^2} [\sigma_1^2 - \sigma_0^2 - (\theta_1 - \theta_0)^2]}$$

$$E(n|\theta = \theta_1, \sigma^2 = \sigma_1^2) = \frac{2[\beta \ln B + (1 - \alpha) \ln A]}{\ln\left(\frac{\sigma_0^2}{\sigma_1^2}\right) + \frac{1}{\sigma_0^2} [\sigma_1^2 - \sigma_0^2 + (\theta_1 - \theta_0)^2]}$$

where

$$A = \frac{1 - \beta}{\alpha}, \quad B = \frac{\beta}{1 - \alpha}$$

Decision Procedure. The decision procedure is as follows:

(1) Compute

$$C = \frac{(\theta_1 - \theta_0)^2}{\sigma_1^2 - \sigma_0^2} + \ln \frac{\sigma_1^2}{\sigma_0^2}, \quad D = \frac{\sigma_1^2 - \sigma_0^2}{\sigma_0^2 \sigma_1^2}, \quad E = \frac{\theta_1 \sigma_0^2 - \theta_0 \sigma_1^2}{\sigma_1^2 - \sigma_0^2}$$

(2) Compute

$$a_m = 2D \ln B + CDm \quad \text{for } m = 1, 2, 3, \dots$$

$$b_m = 2D \ln A + CDm \quad \text{for } m = 1, 2, 3, \dots$$

Random samples of maintenance-task times  $X_1, X_2, \dots$  are obtained as long as

$$a_m < \sum_{i=1}^m (\ln X_i - E)^2 < b_m$$

A decision is made the first time the above inequality is violated.

Thus:

$$\text{Accept } H_0 \text{ after } m \text{ observations if } \sum_{i=1}^m (\ln X_i - E)^2 \leq a_m$$

$$\text{Reject } H_0 \text{ after } m \text{ observations if } \sum_{i=1}^m (\ln X_i - E)^2 \geq b_m$$

Continue testing otherwise.

Since the acceptance and rejection boundaries are linear with  $m$ , the number of sample observations, a graphical procedure is easily established (e.g., in Figure 2, Section IV).

Discussion. The test hypothesis and the use of sequential sampling corresponds to Test Method 1 of MIL-STD-471. However, the latter is based on converting a mean and percentile specification to two percentile specifications and then employing a binomial-type sequential test. This test makes direct use of the lognormal assumption and therefore should have better efficiency in terms of sample size for the comparable MIL-STD-471 test.

The development of the decision criterion is a direct result of application of Wald's theory of the sequential probability-ratio test.

Example. Consider the example of Test Number 5, where specified mean-percentile values led to the hypotheses

$$H_0: \theta = -1.128, \sigma^2 = 0.869 \text{ and } H_1: \theta = -0.715, \sigma^2 = 0.984$$

From the sample-size equations, it is found that  $E(n|\theta = \theta_0, \sigma^2 = \sigma_0^2) = 29$ , and  $E(n|\theta = \theta_1, \sigma^2 = \sigma_1^2) = 26$ . This corresponds to a fixed sample size of 56.

### 7.3.7.3 Test Number 7: Sequential Test on the Mean, Median, or Percentile of a Lognormal Distribution

Description of Test. A sequential test that is based on a conversion of an hypothesized lognormal parameter to an hypothesis on  $\theta = E(\ln X)$ . For a mean or percentile specification, a prior estimate of  $\sigma^2$  is required for the conversion. For a median specification,  $\sigma^2$  need not be estimated. Since the test is based on the asymptotic sufficiency of maximum-likelihood estimators, a minimum sample size should be specified to approximate the asymptotic property.

Underlying Assumptions. Maintenance times can be adequately described by a lognormal distribution. Simple random sampling is performed, and the minimum sample size is large enough so that  $s^2$  is a good estimate of  $\sigma^2$ .

Hypotheses. (The prior estimate of  $\sigma^2$  is denoted by  $\tilde{\sigma}^2$ ).

Mean Specification  $H_0: \mu = \mu_0 \rightarrow \theta = \theta_0 = \ln \mu_0 - \tilde{\sigma}^2/2$

$$H_1: \mu = \mu_1 \rightarrow \theta = \theta_1 = \ln \mu_1 - \tilde{\sigma}^2/2$$

Percentile Specification

$$H_0: X_p = T_0 \rightarrow \theta = \theta_0 = \ln T_0 - Z_p \tilde{\sigma}$$

$$H_1: X_p = T_1 \rightarrow \theta = \theta_1 = \ln T_1 - Z_p \tilde{\sigma}$$

Median Specification

$$H_0: \tilde{M} = M_0 \rightarrow \theta = \theta_0 = \ln M_0$$

$$H_1: \tilde{M} = M_1 \rightarrow \theta = \theta_1 = \ln M_1$$

Sample Size. The sample size is a random variable and is therefore usually evaluated in terms of expected values. Because of the asymptotic nature of this test, however, a minimum

number of samples should be specified. Values of  $n$  greater than 25 should prove satisfactory (see Discussion section). The symbol  $n^*$  is used to denote the specified minimum sample size.

Decision Procedure. Random samples of maintenance-task times  $X_1, X_2, \dots, X_{n^*}, X_{n^*+1}, X_{n^*+2}, \dots$ , are taken as long as the following inequality holds:

$$a_m = \frac{s_m^2}{\theta_1 - \theta_0} \ln B + m \frac{\theta_0 + \theta_1}{2} < \sum_{i=1}^m \ln X_i < \frac{s_m^2}{\theta_1 - \theta_0} \ln A + m \frac{\theta_0 + \theta_1}{2} \\ = b_m, \quad (m \geq n^*)$$

where  $f = \frac{1-\beta}{\alpha}$ ,  $B = \frac{\beta}{1-\alpha}$ ,  $s_m^2 = \frac{1}{m} \left[ \sum_{i=1}^m (\ln X_i)^2 - m \hat{\theta}_m^2 \right]$ , and

$$\hat{\theta}_m = \frac{1}{m} \sum_{i=1}^m \ln X_i.$$

Accept  $H_0$  if  $\sum_{i=1}^m \ln X_i \leq a_m$  for some  $m \geq n^*$

Reject  $H_0$  if  $\sum_{i=1}^m \ln X_i \geq b_m$  for some  $m \geq n^*$

Discussion. Comparison of the decision criterion of this test with a sequential test on the median-- $\sigma^2$  known reveals that the only difference is that  $\sigma^2$  is replaced by  $s^2$  in the accept/reject boundaries. The basis for this substitution is discussed in Reference 4, pp. 54-56.

To indicate the acceptability of this procedure, the sequential test was simulated on a computer. The test hypotheses were  $H_0: M = 30$  minutes and  $H_1: M = 45$  minutes at  $\alpha = \beta = 0.10$ . The minimum sample size before a decision could be made was set at 20. Lognormal random numbers with medians of 30, 36.75, and 45 minutes, and with  $\sigma^2 = 1.44$ , were generated. For 250 simulated tests at each median value, the theoretical and observed results are as follows:

Median	Probability of Acceptance		Expected Sample Size	
	Theoretical	Observed	Theoretical	Observed
30	0.90	0.896	30.79	34.8
36.75	0.50	0.464	42.29	51.4
45	0.10	0.060	30.79	39.1

For the three cases examined, the observed acceptance probabilities are slightly under the theoretical. Because of the minimum-sample-size restriction, the average number of observations is somewhat higher than theoretically expected.

#### 7.3.7.4 Test Number 8: Sequential Test of Critical Maintenance Time, Lognormal Distribution, $\sigma^2$ Unknown

Description of Test. A sequential test of the proportion of a normal population exceeding a given constant. The assumption of a lognormal distribution of maintenance times permits direct application to testing the percentile value of a critical maintenance time. This test is an approximation of the sequential t test (see Discussion section).

Underlying Assumptions. Maintenance times can be adequately described by a lognormal distribution. Simple random sampling is performed. The decision criterion is based on the approximate normality of the Statistic  $\bar{X} + ks$  when samples are taken from a normal distribution.

Hypothesis.  $H_0: P[X > T] = p_0$  or  $T = X_{p_0}$

$H_1: P[X > T] = p_1$  or  $T = X_{p_1}$

where  $T$  is a specified critical maintenance time.

Sample Size. Since  $n$  is a random variable in a sequential test, no sample size is specified.

Decision Procedure. The decision procedure is as follows:

(1) Compute  $A = \ln \frac{1-\beta}{\alpha}$ ,  $B = \ln \frac{\beta}{1-\alpha}$

(2) Compute the acceptance boundary

$$a_m = \frac{\sqrt{1 - C_m D_m} - 1}{C_m}$$

where

$$C_m = \frac{A}{(m-1)(Z_{p_0} - Z_{p_1})}, \quad D_m = \frac{2A}{m(Z_{p_0} - Z_{p_1})} - (Z_{p_0} + Z_{p_1})$$

(3) Compute the rejection boundary

$$b_m = \frac{\sqrt{1 - E_m F_m} - 1}{F_m}$$

where

$$E_m = \frac{B}{(m-1)(Z_{p_0} - Z_{p_1})}, F_m = \frac{2B}{m(Z_{p_0} - Z_{p_1})} - (Z_{p_0} + Z_{p_1})$$

(4) After each observation, compute

$$T_m = \frac{\sum_{i=1}^m \ln X_i}{s_m}, m = 2, 3, \dots$$

where

$$s_m^2 = \frac{1}{m-1} \left[ \sum_{i=1}^m (\ln X_i)^2 - \frac{\left( \sum_{i=1}^m \ln X_i \right)^2}{m} \right]$$

(5) If  $k_m \leq a_m$ ,  $H_0$  is accepted. If  $k_m \geq b_m$ ,  $H_0$  is rejected. If  $a_m < k_m < b_m$ , another maintenance task is sampled.

**Discussion.** This test is the sequential analog of the fixed-sample test for critical maintenance time and, accordingly, should require smaller sample sizes on the average, except possibly for cases in which  $X_{p_0} < T < X_{p_1}$ . The test described here is an ap-

proximation to the WAGR sequential t test (see reference 5), which requires the use of tables of the noncentral t distribution. The derivation of the test is given in Reference 2, pp. 83-85.

### 7.3.8 Nonparametric Tests

#### 7.3.8.1 Introduction

Nonparametric tests have the desirable characteristic that it is not necessary to assume an underlying distribution with regard to the maintenance-time random variable of interest. However, they generally require greater sample sizes than corresponding parametric or nondistribution-free tests.

Five different nonparametric maintainability-demonstration tests are presented in this subsection (the first four are fixed-sample-size tests):

- (1) Test of median or percentile
- (2) Test of critical maintenance time
- (3) Test of two critical maintenance times

- (4) Test of specific lognormal distribution
- (5) Sequential test of critical maintenance time

### 7.3.8.2 Test Number 9 - Test of Median or Percentile

Description of Test. A nonparametric test of a median or percentile. Two specific sets of hypotheses are considered. One is for the case in which a desirable median or percentile is specified such that there should be a high probability of acceptance if the actual maintenance-time distribution conforms. This is the usual null hypothesis with an associated  $\alpha$  risk. The other case offers the consumer protection against accepting a system that has the median or percentile equal to a specified value representing an undesirable maintenance level. The test statistic for either case is based on the number of observations exceeding the specified time.

Underlying Assumption. Since the test is nonparametric, it is not necessary to make an assumption concerning the distribution of maintenance times.

#### Hypotheses.

#### Test I

#### Test II

$$H_0: X_p = T_0$$

$$H_0: X_p < T_1$$

$$H_1: X_p > T_0$$

$$H_1: X_p = T_1$$

For test I,  $T_0$  represents a desirable value for the  $(1-p)^{th}$  percentile; e.g.,  $H_0$  might be: median = 30 minutes and therefore  $H_1$  is the composite alternative, median > 30 minutes. For test II,  $T_1$  represents an undesirable value for the  $(1-p)^{th}$  percentile; e.g.,  $H_1$  might be: median = 45 minutes and  $H_0$  is therefore median < 45 minutes. Note that the hypothesis  $H_1: X_p = T_1$  of test II is the same form as the  $H_0$  hypothesis of test I.

Sample Size. Since both test I and test II contain a composite hypothesis and therefore only one specified risk, there is no sample-size restriction in the usual sense. If  $r$  is the number of maintenance times exceeding the specified value and  $c$  is the acceptance number, an accept decision is made if  $r \leq c$  when  $n$  maintenance times are observed. For a given  $c$  value ( $c = 0, 1, 2, \dots$ ),  $n$  is found from the following equations, which employ the binomial distribution:

$$\text{Test I: } \sum_{r=0}^c \binom{n}{r} p^r (1-p)^{n-r} \geq 1-\alpha$$

$$\text{Test II: } \sum_{r=0}^c \binom{n}{r} p^r (1-p)^{n-r} \leq \beta$$

Table XXXIV presents the sampling plans (sample size  $n$  and acceptance number  $c$ ) for various  $p$ ,  $\alpha$ , and  $\beta$  values for  $c = 0, 1, 2, 3, 4$ , and  $5$ .

TABLE XXXIV

SAMPLE SIZES FOR TESTS I AND II FOR VARIOUS  
PERCENTILES AND RISKS FOR  $c = 0$  THROUGH  $5$

c	Risk	Percentile							
		50th (Median)		80th		90th		95th	
		I	II	I	II	I	II	I	II
0	.20	*	3	1	8	2	16	4	32
	.10	*	4	*	11	1	22	2	45
	.05	*	5	*	14	*	29	1	60
1	.20	*	5	4	14	8	29	16	60
	.10	*	7	2	18	5	38	10	78
	.05	*	8	2	22	3	46	7	94
2	.20	3	8	7	21	15	42	30	86
	.10	*	9	6	25	11	52	22	110
	.05	*	11	4	30	8	62	16	130
3	.20	5	10	11	27	23	54	46	110
	.10	4	12	9	32	18	66	35	140
	.05	*	13	7	37	14	76	28	160
4	.20	6	12	15	33	31	66	62	140
	.10	5	14	13	38	25	78	49	160
	.05	5	16	10	44	20	90	40	190
5	.20	8	15	20	39	39	78	78	160
	.10	7	17	16	45	32	92	62	190
	.05	6	18	14	50	27	110	52	220
* Risk requirement cannot be satisfied.									

In general, the plan with the higher  $c$  number offers better protection against accepting a poor product (for test I) and rejecting a good product (test II); therefore,  $c$  should be made as large as possible, consistent with constraints on sample size.

Decision Procedure. A random sample of  $n$  maintenance times  $X_1, X_2, \dots, X_n$  is observed, and a count is taken of the number of such times that exceed the specified time  $T$ . This number is called  $r$ .

For test I,  $H_0$  is accepted if  $r \leq c$  and is rejected otherwise.

For test II,  $H_1$  is accepted if  $r \leq c$  and is rejected otherwise.

Discussion. Test II is equivalent to Plan 4 of MIL-STD-471. Test I is an alternative in which a desirable percentile value is specified. The choice between I and II depends on whether it is better to control the  $\alpha$  or the  $\beta$  risk, which in turn depends on the costs associated with each of the possible wrong decisions. The next test described has a control on both  $\alpha$  and  $\beta$  risks and is therefore recommended as a better alternative because it includes the specified hypotheses of this test.

### 7.3.3.3 Test Number 10 - Test of Critical Maintenance Time

Description of Test. A nonparametric test of a critical maintenance time and associated percentile value. An example is the following set of hypotheses:  $H_0$  -- 30 minutes is the median (50th percentile); and  $H_1$  -- 30 minutes is the 25th percentile. In this test both the null and alternative hypotheses refer to a fixed time and the percentile value varies. In the preceding test the percentile value remains fixed and the time varies.

As in the preceding test, the number of maintenance-time observations exceeding the critical time is compared with an acceptance number  $c$  to determine acceptance or rejection of the null hypothesis.

Underlying Assumptions. No specific assumption is necessary concerning the distribution of maintenance time. In the development of the equations for determining the decision criterion and sample size, the normal or Poisson approximation to the binomial distribution is used.

Hypothesis.

$$\begin{aligned} H_0: T &= X_{p_0} \\ H_1: T &= X_{p_1} \end{aligned} \quad (p_1 > p_0)$$

Sample Size, n, and Acceptance Number, c. The normal approximation to the binomial distribution is employed to find n and c when  $p_0$  is not a small value. Otherwise, the Poisson approximation is employed. The equations for n and c are as follows:

For  $0.20 \leq p_0 \leq 0.80$

$$n = \left[ \frac{z_\beta \sqrt{p_1 q_1} + z_\alpha \sqrt{p_0 q_0}}{p_1 - p_0} \right]^2 \quad (\text{Use next higher integer value.})$$

$$c = n \left[ \frac{z_\beta p_0 \sqrt{p_1 q_1} + z_\alpha p_1 \sqrt{p_0 q_0}}{z_\alpha \sqrt{p_0 q_0} + z_\beta \sqrt{p_1 q_1}} \right] \quad (\text{Use next lower integer value.})$$

For  $p_0 < 0.20$

For this case n and c can be found from the following two equations:

$$\sum_{r=0}^c \frac{e^{-np_0} (np_0)^r}{r!} \geq 1 - \alpha$$

$$\sum_{r=0}^c \frac{e^{-np_1} (np_1)^r}{r!} \leq \beta$$

Table XXXV provides sampling plans for various  $\alpha$  and  $\beta$  risks and ratios  $p_1/p_0$  when  $p_0 < 0.20$ .

Decision Procedure. Random samples of maintenance times are taken, yielding n observations  $X_1, X_2, \dots, X_n$ . The number of such observations exceeding the specified time T is counted. This number is called r.

Accept  $H_0$  if  $r \leq c$ .

Reject  $H_0$  if  $r > c$ .

Discussion. This plan corresponds most closely to test method 4 of MIL-STD-471 except that both an  $\alpha$  and  $\beta$  risk are specified. It is preferred over the preceding plan because it provides known protection to both the producer and consumer.

TABLE XXXV  
 SAMPLING PLANS FOR SPECIFIED  $p_0$ ,  $p_1$ ,  $\alpha$ , and  $\beta$   
 WHEN  $p_0$  IS SMALL (e.g.,  $p_0 < 0.20$ )

$k = \frac{p_1}{p_0}$	$\alpha = 0.05$						$\alpha = 0.10$						$\alpha = 0.20$					
	$\beta = 0.05$		$\beta = 0.10$		$\beta = 0.20$		$\beta = 0.05$		$\beta = 0.10$		$\beta = 0.20$		$\beta = 0.05$		$\beta = 0.10$		$\beta = 0.20$	
	c	D	c	D	c	D	c	D	c	D	c	D	c	D	c	D	c	D
1.5	66	54.1	54	43.4	39	30.2	51	43.0	40	33.0	29	23.2	36	31.8	27	23.5	17	14.4
2	22	15.7	18	12.4	14	9.25	17	12.8	14	10.3	10	7.02	12	9.91	9	7.29	6	4.73
2.5	13	8.46	10	6.17	8	4.70	10	7.02	8	5.43	6	3.90	7	5.58	5	3.84	3	2.30
3	9	5.43	7	3.98	6	3.29	7	4.66	5	3.15	4	2.43	4	3.09	3	2.30	2	1.54
4	6	3.29	5	2.61	4	1.97	4	2.43	3	1.75	2	1.10	3	2.30	2	1.54	1	0.824
5	4	1.97	3	1.37	3	1.37	3	1.75	2	1.10	2	1.10	2	1.54	1	0.824	1	0.824
10	2	0.818	2	0.818	1	0.353	1	0.532	1	0.532	1	0.532	1	0.824	1	0.824	0	0.227

To find the sample size  $n$ , for given  $p_0$ ,  $p_1$ ,  $\alpha$ , and  $\beta$ , divide the appropriate  $D$  value by  $p_0$  and use the greatest integer less than the quotient. Example:  $p_0 = 0.05$ ,  $p_1 = 0.20$ ,  $\alpha = 0.10$ ,  $\beta = 0.05$ , and  $k = \frac{0.20}{0.05} = 4$ . Then  $n = \frac{2.43}{0.05} = 48.6$ . The acceptance number is  $c = 4$ .

An example of developing the decision criterion for the tail-percentile hypothesis is provided in Table XXXV. For an example of a median specification, assume that the following hypotheses exist, with  $\alpha = \beta = 0.10$ :

$$H_0: 30 \text{ minutes} = X_{0.50} = \text{median}$$

$$H_1: 30 \text{ minutes} = X_{0.75} = 25^{\text{th}} \text{ percentile}$$

Then  $Z_\alpha = Z_\beta = 1.28$  and

$$n = (1.28)^2 \left[ \frac{\sqrt{(.75)(.25)} + \sqrt{(.50)(.50)}}{(.25)} \right]^2 \approx 23$$

and

$$c = 23 \left[ \frac{.5 \sqrt{(.75 \times .25)} + .75 \sqrt{(.50 \times .50)}}{\sqrt{(.50)(.50)} + \sqrt{(.75)(.25)}} \right] \approx 14$$

The actual probabilities corresponding to  $1-\alpha = 0.90$  at  $p_0 = 0.50$ , and to  $\beta = 0.10$  at  $p_1 = 0.75$ , from tables of the binomial distribution, are 0.895 and 0.096, respectively -- an excellent agreement.

#### 7.3.8.4 Test Number 11 - Test of Two Critical Maintenance Times

Description of Test. A nonparametric test in which two critical maintenance times are specified and separate tests are applied to each. The two critical maintenance times will generally be a median and  $M_{\text{max}}$  value. An overall accept decision is made only if both individual tests are passed. As in tests 9 and 10, the number of observed maintenance times exceeding specified values is the statistic used to make the accept/reject decision.

Underlying Assumptions. No specific assumptions concerning the distribution of maintenance times are necessary.

Hypotheses.  $H_0: T = X_{p_0}, t = Y_{q_0}$

$$(T < t, p_0 < p_1, q_0 < q_1, p_1 > q_1, \\ i = 0 \text{ or } 1)$$

$$H_1: T = X_{p_1}, t = X_{q_1}$$

An example of the hypotheses is as follows:

- $H_0$ : 30 minutes = median ( $X_{0.50}$ ),  
 60 minutes = 95th percentile ( $X_{0.05}$ )
- $H_1$ : 30 minutes = 35th percentile ( $X_{0.65}$ ),  
 60 minutes = 90th percentile ( $X_{0.10}$ )

Thus

$$T = 30, t = 60, p_0 = 0.50, p_1 = 0.65, q_0 = 0.05, q_1 = 0.10.$$

Sample Size and Decision Procedure. For given  $\alpha$  and  $\beta$  risks, the determination of appropriate sample size and critical values is a relatively complex calculation. Table XXXVI and Table XXXVII present various sampling plans for  $\alpha = 0.10$ , and  $0.20$ , values of  $n$  from 20 to 80 in steps of 10, and various combinations of  $p_0$ ,  $q_0$ ,  $p_1$ ,  $q_1$ .

In Table XXXVI, the plans are such that the median and 90th or 95th percentile are specified in the  $H_0$  hypothesis. The following four combinations are included:

<u><math>H_0</math> Specification</u>			<u><math>H_1</math> Specification</u>		
(1)	(2)	(3)	(4)	(5)	
0.50	0.10		0.65	0.15	} Median and 90th percentile for $H_0$
0.50	0.10		0.75	0.20	
0.50	0.05		0.65	0.10	} Median and 95th percentile for $H_0$
0.50	0.05		0.75	0.15	

In Table XXXVII, the plans are such that the median and 90th or 95th percentile are specified in the  $H_1$  hypothesis. The following four combinations are included:

<u><math>H_0</math> Specification</u>			<u><math>H_1</math> Specification</u>		
(1)	(2)	(3)	(4)	(5)	
0.40	0.05		0.50	0.10	} Median and 90th percentile for $H_1$
0.40	0.025		0.50	0.10	
0.30	0.025		0.50	0.05	} Median and 95th percentile for $H_1$
0.30	0.01		0.50	0.05	

TABLE XXXVI  
SAMPLING PLANS FOR  $H_0$  SPECIFICATION

Median and 90th Percentile													
For $H_0$ : $p_0 = 0.50$ , $q_0 = 0.10$ For $H_1$ : $p_1 = 0.65$ , $q_1 = 0.15$							For $H_0$ : $p_0 = 0.50$ , $q_0 = 0.10$ For $H_1$ : $p_1 = 0.75$ , $q_1 = 0.20$						
n	$\alpha = 0.10$			$\alpha = 0.20$			n	$\alpha = 0.10$			$\alpha = 0.20$		
	C	D	$\beta$	C	D	$\beta$		C	D	$\beta$	C	D	$\beta$
20	13	4	0.52	13	3	0.43	20	13	4	0.17	13	3	0.00
30	20	5	0.50	20	4	0.38	30	20	5	0.12	20	4	0.08
40	30	6	0.58	24	6	0.23	40	30	6	0.20	24	6	0.01
50	31	8	0.30	29	7	0.13	50	31	8	0.02	29	7	0.00
60	37	9	0.24	35	8	0.11	60	37	9	0.01	35	8	0.00
70	44	10	0.25	41	9	0.09	70	44	10	0.01	41	9	0.00
80	47	12	0.11	48	10	0.10	80	47	12	0.00	48	10	0.00

Median and 95th Percentile													
For $H_0$ : $p_0 = 0.50$ , $q_0 = 0.05$ For $H_1$ : $p_1 = 0.65$ , $q_1 = 0.10$							For $H_0$ : $p_0 = 0.50$ , $q_0 = 0.05$ For $H_1$ : $p_1 = 0.75$ , $q_1 = 0.15$						
n	$\alpha = 0.10$			$\alpha = 0.20$			n	$\alpha = 0.10$			$\alpha = 0.20$		
	C	D	$\beta$	C	D	$\beta$		C	D	$\beta$	C	D	$\beta$
20	14	2	0.54	12	2	0.31	20	14	2	0.19	12	2	0.06
30	20	3	0.45	21	2	0.35	30	20	3	0.09	21	2	0.07
40	25	4	0.31	24	3	0.16	40	25	4	0.02	24	3	0.01
50	30	5	0.20	29	4	0.11	50	30	5	0.01	29	4	0.00
60	37	5	0.18	37	4	0.12	60	37	5	0.00	37	4	0.00
70	42	6	0.13	41	5	0.07	70	42	6	0.00	41	5	0.00
80	47	7	0.09	45	6	0.03	80	47	7	0.00	45	6	0.00

TABLE XXXVII  
SAMPLING PLANS FOR  $H_1$  SPECIFICATION

Median and 90th Percentile													
For $H_0$ : $p_0 = 0.40$ , $q_0 = 0.05$ For $H_1$ : $p_1 = 0.50$ , $q_1 = 0.10$							For $H_0$ : $p_0 = 0.40$ , $q_0 = 0.025$ For $H_1$ : $p_1 = 0.50$ , $q_1 = 0.10$						
n	$\alpha = 0.10$			$\alpha = 0.20$			n	$\alpha = 0.10$			$\alpha = 0.20$		
	C	D	$\beta$	C	D	$\beta$		C	D	$\beta$	C	D	$\beta$
20	12	2	0.61	10	2	0.45	20	13	1	0.38	10	1	0.28
30	17	3	0.56	18	2	0.39	30	16	2	0.33	17	1	0.17
40	21	4	0.48	20	3	0.29	40	22	2	0.20	19	2	0.14
50	25	5	0.40	24	4	0.24	50	25	3	0.18	25	2	0.08
60	31	5	0.33	31	4	0.21	60	31	3	0.11	32	2	0.05
70	35	6	0.29	34	5	0.17	70	40	3	0.07	33	3	0.04
80	39	7	0.26	37	6	0.13	80	30	4	0.06	38	3	0.02

Median and 95th Percentile													
For $H_0$ : $p_0 = 0.30$ , $q_0 = 0.025$ For $H_1$ : $p_1 = 0.50$ , $q_1 = 0.05$							For $H_0$ : $p_0 = 0.30$ , $q_0 = 0.01$ For $H_1$ : $p_1 = 0.50$ , $q_1 = 0.05$						
n	$\alpha = 0.10$			$\alpha = 0.20$			n	$\alpha = 0.10$			$\alpha = 0.20$		
	C	D	$\beta$	C	D	$\beta$		C	D	$\beta$	C	D	$\beta$
20	11	1	0.58	8	1	0.21	20	9	1	0.34	10	0	0.24
30	13	2	0.26	13	1	0.20	30	13	1	0.20	11	1	0.07
40	18	2	0.25	15	2	0.06	40	17	1	0.11	15	1	0.04
50	20	3	0.09	19	2	0.04	50	22	1	0.09	19	1	0.03
60	24	3	0.06	26	2	0.10	60	23	2	0.03	23	1	0.02
70	32	3	0.18	25	3	0.01	70	27	2	0.02	27	1	0.01
80	31	4	0.02	30	3	0.01	80	30	2	0.01	33	1	0.01

The decision procedure is based on two critical values, C and D. An accept decision is made only if C or fewer observed maintenance times are greater than the specified median value (T) and D or fewer observed maintenance times are greater than t, the time corresponding to the  $M_{\max}$  percentile.

For a given  $\alpha$ , a plan is identified by the triplet (n, C, D). Tables XXXVI and XXXVII present, for each n, the smallest value of C and the corresponding D value that satisfy the risk. The appropriate plan to use is then determined from evaluation of the  $\beta$  risks shown in Tables XXXVI and XXXVII for each triplet. In general,  $\beta$  will decrease as C or D decreases with n fixed, or  $\beta$  will decrease as n increases for fixed C.

Since n was limited to only seven values and C limited to only the smallest possible value for a given  $\alpha$  risk, a desired  $\beta$  risk may not be obtainable from the tables (see Discussion section).

Discussion. This plan corresponds closely to Test Method No. 4, MIL-STD-471 (as revised by Change Notice 1). In method 4, however, only undesirable (large rejection probability) median and  $M_{\max}$  values are specified and no consideration is given to the producer's risk. In this test, both producer and consumer risks can be controlled.

The details for determining producer and consumer risks for any given triplet (n, C, D) are developed in Reference 6. The equations are as follows:

Producer's Risk =

$$1 - \sum_{j=0}^C \binom{n}{j} p_0^j (1-p_0)^{n-j} \sum_{k=0}^m \binom{j}{k} \left(\frac{q_0}{p_0}\right)^k \left(1 - \frac{q_0}{p_0}\right)^{j-k}$$

$$= 1 - \sum_{j=0}^C \sum_{k=0}^m \frac{n!}{(n-j)! k! (j-k)!} (1-p_0)^{n-j} q_0^k (p_0 - q_0)^{j-k}$$

Consumer's Risk =

$$\sum_{j=0}^C \binom{n}{j} p_1^j (1-p_1)^{n-j} \sum_{k=0}^m \binom{j}{k} \left(\frac{q_1}{p_1}\right)^k \left(1 - \frac{q_1}{p_1}\right)^{j-k}$$

$$= \sum_{j=0}^C \sum_{k=0}^m \frac{n!}{(n-j)! k! (j-k)!} (1-p_1)^{n-j} q_1^k (p_1 - q_1)^{j-k}$$

where  $m$  = the smaller of the two values  $j$  and  $D$ .

The O.C. curve of a plan based on  $n$ ,  $C$ , and  $D$  can be developed from the above equation for consumer's risk by varying the  $p$  and  $q$  values. Because of the wide variety of alternatives generated by the triplets  $(n, C, D)$ , such evaluation may be warranted in cases where careful control of risk and accepting size is desired.

Examples. Two examples are given:

Example 1 - Assume that it is desired to accept with a 90-percent probability an equipment with a median value of 20 minutes and 95th percentile of 45 minutes. If, however, 20 minutes is the 25th percentile and 45 minutes is the 85th percentile, only a 10-percent acceptance probability is desired. In terms of the standard hypothesis for this test,

$$H_0: 20 \text{ minutes} = X_{0.50}, 45 \text{ minutes} = X_{0.05}$$

$$H_1: 20 \text{ minutes} = X_{0.75}, 45 \text{ minutes} = X_{0.15}$$

with  $\alpha = \beta = 0.10$ . We thus have  $p_0 = 0.50$ ,  $q_0 = 0.05$ ,  $p_1 = 0.75$ ,  $q_1 = 0.15$ .

Since the median and the maximum percentile are specified in  $H_0$ , Table XXXVI is appropriate. For the specified  $p$ ,  $q$  values, the plan closest to meeting the  $\alpha$  and  $\beta$  risks is

$$n = 30, C = 20, D = 3$$

Thus 30 maintenance times are sampled. The equipment passes the test only if 20 or fewer such times are more than 20 minutes long and 3 or fewer actions took over 45 minutes.

Example 2 - Assume that if the median is 1 hour and the 95th percentile is 2 hours, only a 10-percent acceptance probability is desired. However, there should not be more than a 20-percent rejection probability if 1 hour is the 70th percentile and 2 hours is the 99th percentile. Thus

$$H_0: 1 \text{ hour} = X_{0.30}, 2 \text{ hours} = X_{0.01}$$

$$H_1: 1 \text{ hour} = X_{0.50}, 2 \text{ hours} = X_{0.05}$$

with  $\alpha = 0.20$  and  $\beta = 0.10$ . We thus have  $p_0 = 0.30$ ,  $q_0 = 0.01$ ,  $p_1 = 0.50$ ,  $q_1 = 0.05$ .

Since the median and maximum percentile values are specified in the  $H_1$  hypothesis, Table XXXVII is appropriate. From this table, it is seen that the specified  $\alpha$  and  $\beta$  risks are satisfied if  $n = 30$ ,  $C = 11$ ,  $D = 1$ .

#### 7.3.8.5 Test Number 12 - Test for Specific Lognormal Distribution

Description of Test. A nonparametric test of the hypothesis that maintenance times are lognormally distributed with parameters  $\theta$  and  $\sigma^2$ . The Kolmogorov-Smirnov statistic is used as the test statistic. A single alternative hypothesis is not specified and thus no Beta risk is associated with  $H_1$ .

Underlying Assumptions. The distribution of maintenance times is continuous.

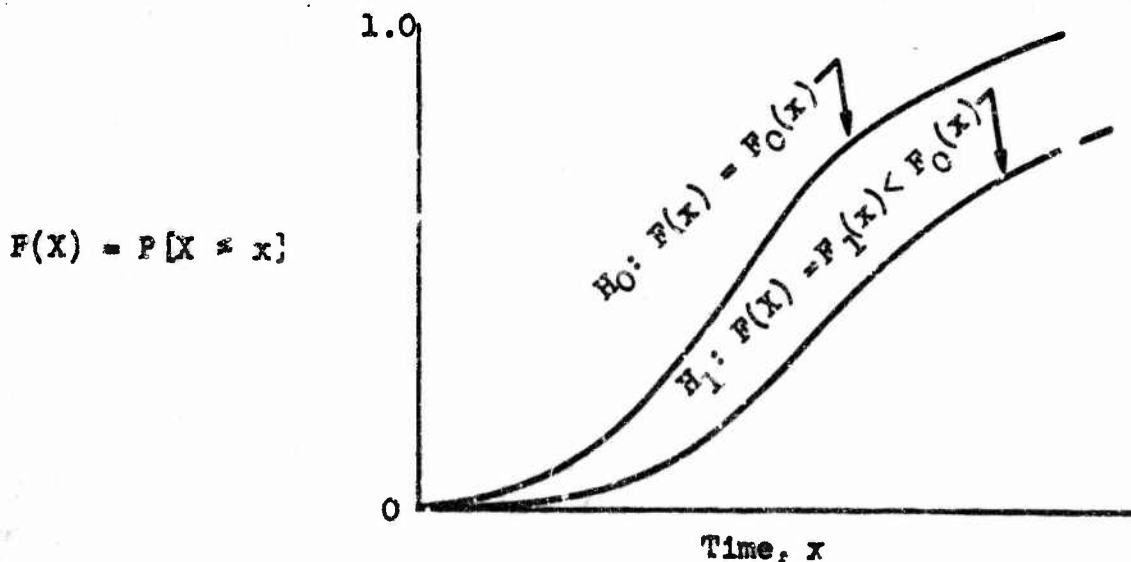
Hypothesis. Let  $F(x)$  be the distribution function of maintenance times, i.e.,  $F(x) = P(X \leq x)$  and  $F_L(x; \theta, \sigma^2)$ , the distribution function of a log-normal with parameters  $\theta$  and  $\sigma^2$ . Then

$$H_0: F(x) = F_0(x) = F_L(x; \theta_0, \sigma_0^2)$$

$$H_1: F(x) = F_1(x) \leq F_0(x) \text{ for all } x$$

Table XXXII of this subsection should be used to find  $\theta_0$  and  $\sigma_0^2$  if another pair of lognormal parameters (e.g., mean and percentile) is specified.

The graphic illustration of the two hypothesized distribution functions is shown below. For any value of  $x$ , the probability of completing a maintenance action is greater under  $H_0$  than under  $H_1$ .



Sample Size. No specific equation for sample size can be given. The greater  $n$  is, however, the more powerful the test will be against any specific alternative hypothesis.

Decision Criterion. A random sample of  $n$  maintenance times is observed. It is necessary to compute the statistic  $D$ , which is defined by the equation

$$D = \max_x \left[ F_L(x; \theta_0, \sigma_0^2) - S(x) \right]$$

where  $F_L(x; \theta_0, \sigma_0^2)$  is the hypothesized lognormal distribution function and  $S(x)$  is the observed distribution function.

To be able to use published tables of the  $D$  statistic, it is necessary first to consider the variable  $Y_1 = -\ln X_1$  where  $X_1$  is the 1<sup>th</sup>-ordered observed maintenance time, i.e.,  $X_1 \leq X_2 \leq X_3 \leq \dots \leq X_{n-1} \leq X_n$ . Then the statistic to be computed is

$$D = \max_y \left[ S(y) - F_N(y; \theta_0, \sigma_0^2) \right]$$

where  $F_N(y; \theta_0, \sigma_0^2)$  is the normal distribution function with mean  $\theta_0$  and variance  $\sigma_0^2$ .

The procedure is as follows:

(1) Let  $Y = -\ln X$  for each of the observed  $n$  times  $X_1, X_2, \dots, X_n$

(2) Order the  $Y$ 's such that  $Y_1 \leq Y_2 \leq \dots \leq Y_{n-1} \leq Y_n$

(3) Compute

$$S(Y_1) = \frac{\text{number of } Y\text{'s less than or equal to } Y_1}{n}$$

(4) Compute for each  $i$  the normal deviate,

$$Z_1 = \frac{-Y_1 - \theta_0}{\sigma_0}$$

- (5) For each  $Z_1$ , use normal probability tables to obtain the normal probability

$$F_N(-Y_1; \theta_0, \sigma_0^2) = F_N(Z_1; 0, 1)$$

which is equal to the probability associated with a normal deviate of  $Z_1$ ,

$$\text{i.e., } F_N(-Y_1, \theta_0, \sigma_0^2) = \int_{-\infty}^{Z_1} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

- (6) Let  $G_N(-Y_1; \theta_0, \sigma_0^2) = 1 - F_N(-Y_1; \theta_0, \sigma_0^2)$

- (7) For each  $Y_1$ , compute

$$\max_{Y_1} [S(Y_1) - G_N(Y_1; \theta_0, \sigma_0^2)]$$

as follows:

For  $i=1$ , compute  $S(Y_1) - G_N(Y_1)$

For  $1 < i \leq n$ , compute

$S(Y_{i-1}) - G_N(Y_{i-1})$  and  $S(Y_i) - G_N(Y_i)$

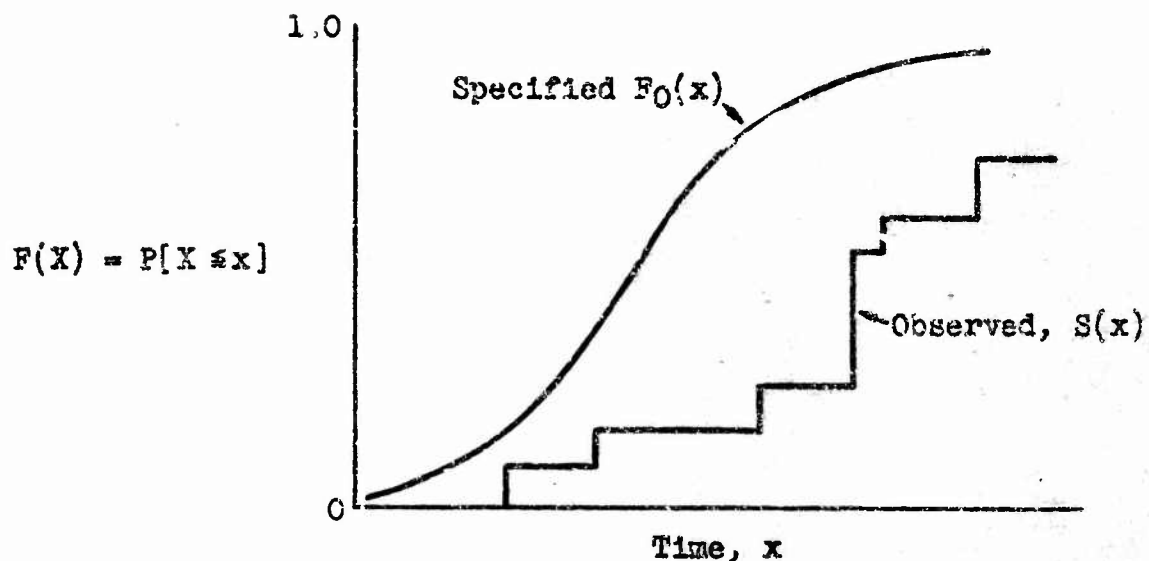
$D$  is then equal to the largest of the above differences.

- (8) For a given risk, refer to a table of the Kolmogorov-Smirnov statistic for the critical  $D$  value,  $D_{\alpha, n}$ .

Accept  $H_0$  if  $D \leq D_{\alpha, n}$ .

Reject  $H_0$  otherwise.

Discussion. This test is a one-sided Kolmogorov-Smirnov test, which is described in Reference 7. From the stated hypotheses of the test, if the observed distribution function is to the right of the theoretical (as specified by  $H_0$ ), e.g.,



then the data would tend to favor  $H_1$  and thus the hypothesis would be rejected. In this case, the statistic  $S(x) - F_0(x)$  is small (in fact, it is negative). So that tabled values of the D statistic can be used, the conversion is made to negative values of the observed maintenance times and  $[1 - F_0(-x)]$  is used with the result that small values of  $[S(x) - F_0(x)]$  correspond to large values of  $[S(-x) - (1 - F_0(-x))]$ .

The conversion to logarithms is first made solely so that it will be possible to use the normal distribution tables rather than have to develop tables of the lognormal-distribution function.

Example. Assume the following:

$$H_0 \text{ is } F(x) = F_L(x; \theta=3, \sigma^2 = 0.25).$$

$$H_1 \text{ is } F(x) < F_L(x; \theta=3, \sigma^2 = 0.25).$$

$\alpha$  is set equal to 0.10.

Assume that a sample of 10 observations yields the following maintenance times (minutes): 15.4, 14.8, 30.1, 35.6, 12.7, 14.8, 24.8, 63.4, 13.0, and 38.4. The table at the top of the following page is developed in accordance with steps 1 through 6.

X	Steps 1 and 2 $Y_1 = -\ln X_1$	Step 3 $S(Y_1)$	Step 4 $Z_1 = \frac{-Y_1 - \theta}{\sigma}$	Step 5 $F_N(Y_1)$	Step 6 $G_N(Y_1) = 1 - F_N(Y_1)$
63.4	-4.14	0.1	2.28	0.99	0.01
38.4	-3.65	0.2	1.30	0.90	0.10
35.6	-3.57	0.3	1.14	0.87	0.13
30.1	-3.40	0.4	0.80	0.79	0.21
24.8	-3.21	0.5	0.42	0.66	0.34
15.4	-2.73	0.6	-0.54	0.30	0.70
14.8	-2.69	0.8	-0.62	0.27	0.73
14.8	-2.69	0.8	-0.62	0.27	0.73
13.0	-2.56	0.9	-0.88	0.19	0.81
12.7	-2.54	1.0	-0.92	0.18	0.82

Then for Step 7, the following are obtained:

<u>i</u>	<u><math>S(Y_{i-1}) - G_n(Y_1)</math></u>	<u><math>S(Y_1) - G_n(Y_1)</math></u>
1	----	0.09
2	0	0.10
3	0.07	0.17
4	0.09	0.19 (max)
5	0.06	0.16
6	-0.20	-0.10
7 } 8 }	-0.13	0.07
9	-0.01	0.09
10	0.08	0.18

Since the maximum difference is 0.19, which is less than the critical value of 0.32 for a sample size of 10, and  $\alpha = 0.10$ , the null hypothesis of a lognormal distribution with  $\theta = 3$  and  $\sigma^2 = 0.25$  cannot be rejected.

#### 7.3.8.6 Test Number 13 - Sequential Test of a Critical Maintenance Time

Description of Test. A nonparametric sequential test of the probability of exceeding a specified maintenance time. Since the

procedure is based solely on the number of maintenance times exceeding the critical time, it is equivalent to a sequential test of a binomial parameter,  $p$ .

Underlying Assumptions. No specific form of the distribution of maintenance times is assumed. Maintenance tasks are sampled through a simple random procedure.

Hypotheses.  $H_0: T = X_{p_0}$  or  $P[X > T] = p_0$

$H_1: T = X_{p_1}$  or  $P[X > T] = p_1$

( $\alpha$  and  $\beta$  are specified)

Sample Size. The sample size of a sequential test is a random variable. Table XXXVIII presents the expected number of observations before a decision for three values of  $p$ .

TABLE XXXVIII

EXPECTED SAMPLE SIZES FOR VALUES OF  $p$

$p$	$E(n p)$
$p_0$	$\frac{(1-\alpha) \ln B + \alpha \ln A}{p_0 \ln C + (1-p_0) \ln D}$
$\frac{\ln D}{\ln D - \ln C}$	$\frac{\ln A \ln B}{\ln C \ln D}$
$p_1$	$\frac{\beta \ln B + (1-\beta) \ln A}{p_1 \ln C + (1-p_1) \ln D}$
Notation: $A = (1-\beta)/\alpha$ , $B = \beta/(1-\alpha)$ $C = (p_1/p_0)$ , $D = [(1-p_1)/(1-p_0)]$	

If desired, an upper bound on the sample size can be established that will have limited effects on the  $\alpha$  and  $\beta$  risks. The suggested procedure is to truncate the test at three times the expected sample size, say at  $m=n^*$ . The expected sample size to use will depend on the assumed value of  $p$  in the expression  $E(n|p)$ .

Generally,  $p_0$  is selected unless prior evidence indicates otherwise; therefore,

$$n^* = 3E(n|p = p_0) = 3 \frac{(1-\alpha)\ln B + \alpha\ln A}{p_0 \ln \frac{p_1}{p_0} + (1-p_0) \ln \frac{1-p_1}{1-p_0}}$$

Decision Procedure. The decision procedure is as follows:

(1) Compute  $A = \frac{1-\beta}{\alpha}$ ,  $B = \frac{\beta}{1-\alpha}$

(2) Compute  $C = \ln \frac{p_1}{p_0}$ ,  $D = \ln \frac{1-p_0}{1-p_1}$

(3) Compute the acceptance boundary

$$a_m = \frac{\ln B}{C + D} + m \frac{D}{C + D} \text{ for } m = 1, 2, \dots$$

(4) Compute the rejection boundary

$$b_m = \frac{\ln A}{C + D} + m \frac{D}{C + D} \text{ for } m = 1, 2, \dots$$

Random samples of maintenance times  $X_1, X_2, \dots$  are then obtained as long as

$$a_m < d_m < b_m$$

where  $d_m$  is the number of maintenance times that exceed  $T$  after  $m$  observations are made. A decision is made the first time the above inequality is violated:

Accept  $H_0$  if for some  $m$ ,  $d_m \leq a_m$

Reject  $H_0$  if for some  $m$ ,  $d_m \geq b_m$

If the truncated sample size (say,  $n^*$ ) is reached and no decision has been made:

$$\text{Accept } H_0 \text{ if } d_{n^*} \leq \frac{(a_{n^*} + b_{n^*})}{2}$$

$$\text{Reject } H_0 \text{ if } d_{n^*} > \frac{(a_{n^*} + b_{n^*})}{2}$$

Discussion. This test and Test Method 1 of MIL-STD-471 are similar in that both are sequential tests of percentile values. However, the MIL-STD-471 plan assumes a mean or  $M_{\max}$  specification under a lognormal assumption, which is then converted to specifications on percentiles.

This test is a standard application of sequential sampling under a binomial assumption. Reference 8, pp. 88-105, presents a detailed discussion of the test's operating characteristics as well as the practical consequences of taking observations in groups.

#### 7.3.8.7 Test Number 14 - Sequential Test of Two Critical Maintenance Times

Description of Test. This test is the sequential counterpart of Test Number 11. Two critical maintenance times are specified, generally a median and  $M_{\max}$  value. The decision criterion is developed from a direct application of the sequential probability-ratio test.

Underlying Assumptions. No specific assumptions concerning the distribution of maintenance times are necessary. Simple random sampling is performed.

##### Hypotheses.

$$H_0: T = 100 \times (1-p_0)\text{percentile} = X_{p_0},$$

$$t = 100 \times (1-q_0)\text{percentile} = X_{q_0}$$

$$H_1: T = 100 \times (1-p_1)\text{percentile} = X_{p_1},$$

$$t = 100 \times (1-q_1)\text{percentile} = X_{q_1}$$

where

$$T < t, p_0 < p_1, q_0 < q_1, p_1 > q_1 \text{ for } i = 0 \text{ or } 1$$

Note that  $p_1$  is the percent of observations greater than  $T$  for  $H_1$ , and  $q_1$  is the percent of observations greater than  $t$  for  $H_1$ . Risks of  $\alpha$  and  $\beta$  are also specified.

Sample Size. The sample size of a sequential test is a random variable. For the cases in which  $H_0$  and  $H_1$  are true, the following are the equations for the expected sample size:

$$E(n|H_0 \text{ true}) = \frac{(1-\alpha)\ln B + \alpha\ln A}{(1-p_0)C + (p_0-q_0)D + q_0E}$$

$$E(n|H_1 \text{ true}) = \frac{\beta\ln B + (1-\beta)\ln A}{(1-p_1)C + (p_1-q_1)D + q_1E}$$

where  $A$ ,  $B$ ,  $C$ ,  $D$ , and  $E$  are as defined below.

Decision Procedure.

(1) Compute  $A = \frac{1-\beta}{\alpha}$ ,  $B = \frac{\beta}{1-\alpha}$

(2) Compute  $C = \ln\left(\frac{1-p_1}{1-p_0}\right)$ ,  $D = \ln\left(\frac{p_1-q_1}{p_0-q_0}\right)$ ,  $E = \ln\left(\frac{q_1}{q_0}\right)$

(3) Compute the acceptance boundary

$$a_m = -\ln B + mE$$

(4) Compute the rejection boundary

$$b_m = -\ln A + mE$$

Random samples of maintenance times  $X_1, X_2, \dots$  are then obtained as long as

$$b_m < (D-C)N_m(T) + (E-D)N_m(t) < a_m$$

where

$N_m(T)$  = number of maintenance times less than or equal to  $T$  after  $m$  observations

$N_m(t)$  = number of maintenance times less than or equal to  $t$  after  $m$  observations

A decision is made the first time the above inequality is not true. Thus

Accept  $H_0$  if for some  $m$ ,  $(D-C)N_m(T) + (E-D)N_m(t) \geq a_m$

Reject  $H_0$  if for some  $m$ ,  $(D-C)N_m(T) + (E-D)N_m(t) \leq b_m$

Discussion. This test is comparable to Test Method 1 of MIL-STD-471 in that both are sequential tests of two percentile values. The MIL-STD-471 test, however, is based on a conversion of a lognormal mean/percentile specification to a two-percentile specification; for each such specification a separate sequential decision criterion is then applied. The MIL-STD-471 method appears to be based on an approximation method for using the same data for the two separate tests in order to satisfy the alpha and beta risks.

In this test, only one sequential decision criterion is applied; it takes into account the dependence between the number of observations less than or equal to  $T$ , the number between  $T$  and  $t$ , and the number greater than  $t$ . The test is therefore exact in the sense that any sequential probability-ratio test is exact.

If desired, Table XXXII can be used to convert a lognormal mean/percentile specification (or any other lognormal parameter pair) to a specification of two critical maintenance times so that this test will be applicable. If the lognormal assumption can be reasonably made, however, the parametric counterpart to this test (Test Number 6) will generally be the more desirable choice.

Since this test does not generally appear in the literature, its derivation is outlined below.

The basic inequality for the sequential probability-ratio test is

$$\frac{\beta}{1-\alpha} = B < \frac{\prod_{i=1}^m f(X_i|H_1)}{\prod_{i=1}^m f(X_i|H_0)} = \frac{p_{1m}}{p_{0m}} < A = \frac{1-\alpha}{\beta}$$

where  $f(X_i|H_j)$  is the density of the  $i^{\text{th}}$  observation under  $H_j$ ,  $j = 0, 1$ .

If  $\frac{p_{1m}}{p_{0m}} \geq A$ , then the data indicate that  $H_1$  is significantly more consistent with the observations than  $H_0$ , with the result that  $H_0$  is rejected. Similarly,  $H_0$  is accepted if  $\frac{p_{1m}}{p_{0m}} \leq B$ . For

this problem the observed maintenance times can be classified as being less than or equal to  $T$  or less than or equal to  $t$ , generating the two random variables  $N_m(T)$  and  $N_m(t)$  as defined above. Then it can be shown that

$$p_{jm} = P[N_m(T) = k_0, N_m(t) = k_1]$$

$$= \frac{m!}{k_0! k_1! (m - k_0 - k_1)!} (1 - p_j)^{k_0} (p_j - q_j)^{k_1 - k_0} q_j^{m - k_1}, \quad j = 0, 1$$

which is a trinomial probability function (the third variable is the number of observations greater than  $t$ , which is always equal to  $m - k_1$ ). When the above expressions for  $p_{0m}$  and  $p_{1m}$  are inserted into the basic inequality, the result after taking logarithms and simplifying is as given above in the Decision Procedure section.

For the expected sample-size equations, the basic formula is

$$E(n|H_j) = \frac{L(H_j) \ln B + [1 - L(H_j)] \ln A}{E_j \left[ \frac{\ln \frac{f(X|H_1)}{f(X|H_0)}}{\frac{f(X|H_1)}{f(X|H_0)}} \right]}$$

where

$$L(H_j) = P[\text{accept } H_0 | H_j \text{ true}] \quad (L(H_0) = 1 - \alpha, \quad L(H_1) = \beta)$$

$E_j$  is the expectation operator under the condition  $H_j$  true

Since  $f(X|H_j)$  in this problem is equal to  $f(N_1(T), N_1(t)|H_j)$ ,

$$\begin{aligned} \text{then} \quad f(N_1(T), N_1(t)|H_j) &= (1 - p_j)^{N_1(T)} (p_j - q_j)^{[N_1(t) - N_1(T)]} \\ &\quad \cdot q_j^{[1 - N_1(t)]} \quad \text{for } j = 0, 1 \end{aligned}$$

where

$$N_1(T) = \begin{cases} 0 & \text{if } X > T \\ 1 & \text{if } X \leq T \end{cases} \quad N_1(t) = \begin{cases} 0 & \text{if } X > t \\ 1 & \text{if } X \leq t \end{cases}$$

Then

$$\begin{aligned} E_j \left[ \frac{\ln \frac{f(X|H_1)}{f(X|H_0)}}{\frac{f(X|H_1)}{f(X|H_0)}} \right] &= (1 - p_j) \ln \left( \frac{1 - p_1}{1 - p_0} \right) + (p_j - q_j) \ln \left( \frac{p_1 - q_1}{p_0 - q_0} \right) + \\ &\quad q_j \ln \left( \frac{q_1}{q_0} \right), \quad \text{for } j = 0, 1 \end{aligned}$$

which is the denominator of the sample-size equation

Example. Assume the following set of hypotheses for a test for which  $\alpha = \beta = 0.10$ .

$H_0$ : 0.25 hrs. = median = 50<sup>th</sup> percentile, 1 hr. = 95<sup>th</sup> percentile

$H_1$ : 0.25 hrs. = 35<sup>th</sup> percentile, 1 hr. = 90<sup>th</sup> percentile

In terms of the notation used, the following is equivalent:

$H_0$ : 0.25 =  $T = X_{0.50}$ , 1.0 =  $t = X_{0.05}$

$H_1$ : 0.25 =  $T = X_{0.65}$ , 1.0 =  $t = X_{0.10}$

Hence

$T = 0.25$ ,  $t = 1.0$ ,  $p_0 = 0.50$ ,  $q_0 = 0.05$ ,  $p_1 = 0.65$ ,  $q_1 = 0.10$

Then

$\ln A = \ln 9 = 2.1972$ ,  $\ln B = \ln(1/9) = -2.1972$

$C = \ln \frac{0.35}{0.50} = -0.3567$ ,  $D = \ln \frac{0.55}{0.45} = 0.2007$

$E = \ln \frac{0.10}{0.05} = 0.69315$

The acceptance boundary is  $a_m = 2.1972 + 0.6932m$

The rejection boundary is  $b_m = -2.1972 + 0.6932m$

The test statistic is

$N_m(T)(D-C) + N_m(t)(E-D) = 0.5567N_m(T) + 0.4925N_m(t)$

where

$N_m(T)$  is the cumulative number of observations less than or equal to  $T = 0.25$  hrs.

$N_m(t)$  is the cumulative number of observations less than or equal to  $t = 1.0$  hrs.

For the expected sample sizes,  $E(n|H_0) \approx 34$  and  $E(n|H_1) \approx 33$ . These expected values correspond to a fixed-sample-size requirement of almost 80 observations (see Table XXXVI).

If the hypotheses of the example of Test Number 6 are converted to two critical maintenance times so that Test Number 14 can be applied, the expected sample sizes are 35 if  $H_0$  is true and 32 if  $H_1$  is true. These values correspond to the expected sample sizes of the Test Number 6 application of 29 and 26, respectively, indicating the greater efficiency of the parametric test.

#### 7.4 GUIDELINES FOR TEST SELECTION - NON-BAYESIAN TESTS

##### 7.4.1 Introduction

Fourteen different non-Bayesian tests that can be used for maintainability demonstration have been presented in Subsection 7.3. In this subsection, guidelines are presented for selecting the test that is appropriate for a particular situation.

##### 7.4.2 Summary of the Fourteen Tests

Generally, the factors associated with the maintainability-demonstration program will restrict the choice of test to one of two alternatives. Table XXXIX summarizes the fourteen tests with respect to ten major factors that are relevant to the choice of method.

##### 7.4.3 Decision Tree for Selecting a Test

Table XL is a decision tree derived from Table XXXIX; it indicates which test will meet requirements on type of sampling, distribution assumption, and parameter specification. Thus, a fixed-sample test of the median based on a lognormal assumption should be based on Test 2.

Several alternatives exist in the tree. For example, both tests 3 and 4 are percentile tests based on a fixed sample size and lognormal assumption. Reference to Table XL or Subsections 7.3.6.4 and 7.3.6.5 reveals that test 3 is a test of a critical percentile ( $p$  is fixed), while test 4 is actually a test of a critical maintenance time ( $T$  is fixed), although both have a percentile specification for the null hypothesis.

The distinction between the nonparametric tests 9 and 10 is similar. Sequential test 6 differs from tests 7 and 8 in that it is based on a joint specification. Tests 7 and 8 differ in the same sense as tests 3 and 4. Test 11 differs from test 12 in that the former is a test of a median and  $M_{\max}$  specification without regard to distributional form, while the latter is a test for a specific lognormal distribution, which is defined by the pair of specified parameters.

TABLE XXIX

## SUMMARY OF FOURTEEN TEST PLANS WITH RESPECT TO TEN MAJOR FACTORS

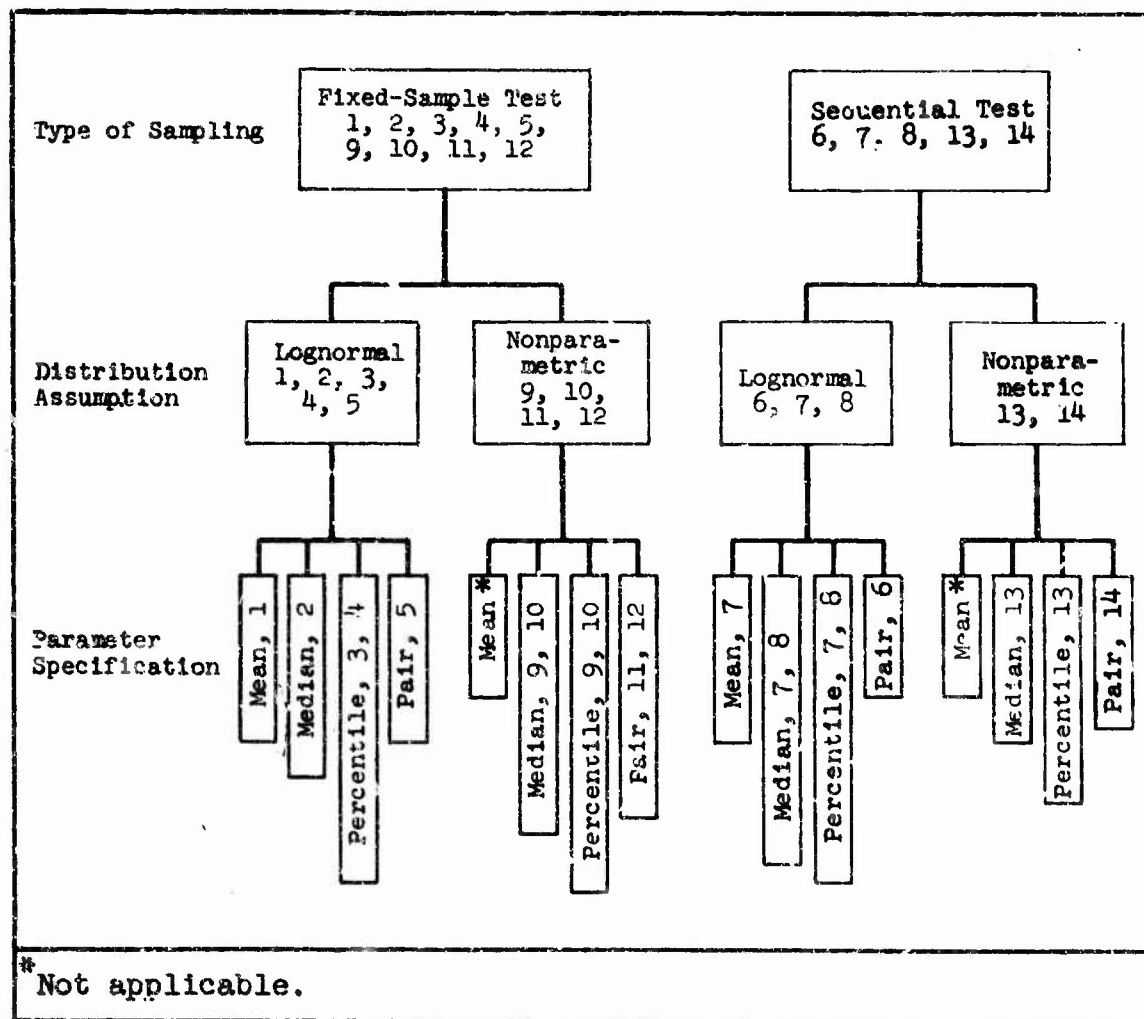
FACTOR										
Test Number	Test Parameters	Null Hypothesis, $H_0$	Alternative Hypothesis, $H_1$	Specified Risk	Fixed or Sequential Test	Sample Size Requirement	Computed Statistics	Distribution Assumption	Other Assumptions	Test Method
1	Mean	Mean = $\mu_0$	Mean = $\mu_1 > \mu_0$	$\alpha$ and $\beta$	Fixed	Equation given	Sample mean and variance of $X$	Lognormal	Control limit unknown. Prior estimates of $\sigma^2$ for obtaining $n$	Test Method 2, mean
2	Median	Median = $\tilde{\mu}_0$	Median = $\tilde{\mu}_1 > \tilde{\mu}_0$	$\alpha$ and $\beta$	Fixed	Equation given	Sample mean and variance of $\ln X$	Lognormal	Prior estimate of $\sigma^2$ for obtaining $n$	Test Method 3
3	Critical Percentile	$p = p_0$	$X_p = T_1 > T_0$	$\alpha$ and $\beta$	Fixed	Equation given ( $n \geq 10$ )	Sample mean and variance of $\ln X$	Lognormal	Approximate normality of sample percentile	Test Method 2, $M_{max}$
4	Critical Maintenance Time	$T = X_{p_0}$	$T = X_{p_1} (p_1 > p_0)$	$\alpha$ and $\beta$	Fixed	Equation given	Sample mean and variance of $\ln X$	Lognormal	Approximate normality of sample percentile	None
5	Pair of Lognormal Parameters	$\sigma^2 = \sigma_0^2$	$\sigma^2 = \sigma_1^2$	$\alpha$ and $\beta$	Fixed	Equation and table given	Sum of squared deviations of $\ln X$	Lognormal	None	None
6	Pair of Lognormal Parameters	$\sigma^2 = \sigma_0^2$	$\sigma^2 = \sigma_1^2$	$\alpha$ and $\beta$	Sequential	Random variable	Sum of squared deviations of $\ln X$	Lognormal	None	Test Method 1
7	Mean, Median, or Percentile	$\theta = \theta_0$	$\theta = \theta_1$	$\alpha$ and $\beta$	Sequential	Random variable; minimum of 25 should be established	Sum and variance of $\ln X$	Lognormal	$\sigma^2$ is used as estimate of $\sigma^2$	None
8	Critical Maintenance Time	$T = X_{p_0}$	$T = X_{p_1}$	$\alpha$ and $\beta$	Sequential	Random variable	Sum and variance of $\ln X$	Lognormal	Approximation to sequential t test	Test Method 1
9	Median or Percentile	$X_p = T_0$	$X_p > T_0$	I - $\alpha$ II - $\beta$	Fixed	Should not be smaller	Number of maintenance times exceeding specified value	None	None	Test Method 4
10	Critical Maintenance Time	$T = X_{p_0}$	$T = X_{p_1} (p_1 > p_0)$	$\alpha$ and $\beta$	Fixed	Equation and table given	Number of maintenance times exceeding specified value	None	Poisson and Normal approximations to binomial	Test Method 3
11	Two Critical Maintenance Times	$T = X_{p_0}$	$T = X_{p_1}$	$\alpha$ and $\beta$	Fixed	Equation and table given	Number of maintenance times exceeding specified value	None	None	Test Method 4
12	Pair of lognormal parameters	$\sigma^2 = \sigma_0^2$	$\sigma^2 = \sigma_1^2$	$\alpha$ and $\beta$	Fixed	Should not be smaller	Observed distribution function of $\ln X$	Continuous	None	None
13	Critical Maintenance Time	$T = X_{p_0}$	$T = X_{p_1} (p_1 > p_0)$	$\alpha$ and $\beta$	Sequential	Random variable	Number of maintenance times exceeding specified value	None	None	None
14	Two Critical Maintenance Times	$T = X_{p_0}$	$T = X_{p_1}$	$\alpha$ and $\beta$	Sequential	Random variable	Number of maintenance times exceeding specified value	None	None	Test Method 1

<sup>a</sup>A minimum value of 25 can be used as a general rule of thumb.

\*A minimum value of 25 can be used as a general rule of thumb.

It is still necessary to choose between fixed or sequential and parametric (lognormal) or nonparametric tests. These alternatives have been discussed in Section IV.

TABLE XL  
DECISION TREE FOR GUIDANCE TO TEST SELECTION



#### 7.4.4 Combinations of Two Tests

In some cases it may be desirable to use two tests -- for example, tests 1 and 4 for a combined mean/percentile test. Where

separate samples are to be used for each individual test, the individual test risks can be determined from overall risks as follows:

Let  $\gamma$  and  $\psi$  be the parameters tested by the two tests, with  $\gamma_0$  and  $\psi_0$  representing desirable levels and  $\gamma_1$  and  $\psi_1$  representing undesirable levels. The overall test risks are defined as follows:

$$\alpha = P[\text{reject if } \gamma = \gamma_0 \text{ and } \psi = \psi_0]$$

$$\beta = P[\text{accept if } \gamma = \gamma_1 \text{ or } \psi = \psi_1]$$

Then, if  $\alpha_1$  and  $\beta_1$  represent the test risks for the  $i^{\text{th}}$  test ( $i=1$  or  $2$ ), under the assumption of independence, the following rule should be observed:

Choose  $\alpha_1$  and  $\alpha_2$  such that

$$(1-\alpha_1)(1-\alpha_2) = 1 - \alpha$$

and

choose  $\beta_1$  and  $\beta_2$  such that

$$\beta_1 + \beta_2 - \beta_1\beta_2 = \beta$$

This rule allows for assigning importance factors for one test over the other. If both tests are considered equally important, the results are as follows:

$$\alpha_1 = \alpha_2 = 1 - (1-\alpha)^{\frac{1}{2}}$$

$$\beta_1 = \beta_2 = 1 - (1-\beta)^{\frac{1}{2}}$$

## 7.5 BAYESIAN TESTS

### 7.5.1 General

Tests for which prior information is incorporated in the decision criterion through use of Bayes' Theorem are called Bayesian tests. If  $\rho$  is the parameter of interest, prior information on  $\rho$  is available in the form of a density function  $g(\rho)$ , and  $n$  sample observations  $X_n = (x_1, x_2, \dots, x_n)$  are made, Bayes' Theorem leads to the following equality:

$$f(\rho|X_n) = \frac{L(X_n|\rho) g(\rho)}{\int_{-\infty}^{\infty} L(X_n|\rho) g(\rho) d\rho}$$

where  $f(\rho|X_n)$  is the posterior density of  $\rho$  after the sample observations  $X_n$  are observed.

$L(X_n|\rho)$  is the likelihood of  $X_n$  given  $\rho$ . Since  $\rho$  is integrated out in the denominator, the above equation can be rewritten as

$$f(\rho|X_n) = \frac{1}{K} L(X_n|\rho) g(\rho)$$

where  $K$  is a proportionality constant. This equation indicates that the posterior distribution of  $\rho$  is equal to the prior distribution modified by the observed test results  $X_n$ .

A Bayesian test can be based on the properties of the posterior distribution. For example, the posterior mean value of the parameter  $\rho$  can be computed from the equation

$$E(\rho|X_n) = \int_{-\infty}^{\infty} \rho f(\rho|X_n) d\rho$$

and the accept/reject decision made according to whether or not  $E(\rho|X_n)$  falls within a desirable region.

There are several advantages of a Bayesian test in maintainability demonstration over so-called classical procedures. If information is available on the maintainability characteristics of an equipment, then a decision procedure employing that information in the form of  $g(\rho)$  plus the additional test information  $X_n$  is obviously more complete than one based solely on  $X_n$ .

Secondly, the practical consequences of a Bayesian test can be important. If the prior density is such that there is high assurance that the equipment is satisfactory, a Bayesian test will generally require relatively little additional testing. On the other hand, if  $g(\rho)$  is unsatisfactory, the product can be accepted only after relatively extensive testing.

Thirdly, a Bayesian test can be constructed to provide a control that may be more pertinent to the needs of the customer. A classical test essentially controls the acceptance probability for specified levels of desirable and undesirable product. Considering the latter, for example, the Beta risk control is defined as

$$P[\text{Accept} | \text{Maintainability is unacceptable}] = \beta$$

A Bayesian test, on the other hand, can be designed to control the maintainability of the accepted equipment by the criterion

$$P[\text{Maintainability is unacceptable} | \text{accept}] = \beta$$

A specific Bayesian maintainability-demonstration test that is based on a control on accepted product maintainability will be presented in this subsection. Because of the newness of this type of test and its broad implications, the condensed format used for describing the nonsequential tests will be replaced by a more detailed description including mathematical derivation of the test criterion.

#### 7.5.2 Basic Assumptions

The basic assumptions are as follows:

- (1) The maintenance-time random variable,  $X$ , has a lognormal distribution with parameters  $\theta$  and  $\sigma^2$  where  $\theta = E(\ln X)$  and  $\sigma^2 = \text{Var}(\ln X)$ .
- (2) The parameter  $\theta$  has a normal prior distribution with mean  $\bar{\theta}$  and variance  $w^2$ .
- (3) The parameter  $\sigma^2$  is known or can be accurately estimated.

Assumption 1 is the usual application of the lognormal distribution for maintenance-time description. Data collected and analyzed by ARINC Research on this and other studies strongly support this assumption. Further support for use of the lognormal distribution for maintenance-time analyses is provided in many other studies involving analysis of observed maintenance times.

Assumption 2 is made for three reasons. First, the use of a normal prior distribution for  $\theta$  allows relatively easy development of a Bayesian test since it leads to a posterior distribution that is also a normal. (Priors that lead to a posterior density of the same form are called conjugate.) Secondly, the normal distribution's symmetrical characteristic and its other well known properties permits the use of easy and known methods of quantifying its parameters either from subjective evaluations or observed

data. Third, the fact that so many real variables of measurement can be adequately described by a normal distribution attests to its versatility. In the absence of data supporting a nonsymmetrical prior, the normal seems as reasonable a choice as any.

The Bayesian test to be described does not depend on a normal prior for its theoretical justification. Nonsymmetrical priors or discrete-type distributions can be employed once suitable modifications are made.

Assumption 3 is made primarily for simplification in development of the procedure. Otherwise, a prior density for  $\sigma^2$  would also have to be employed, leading to quite complex mathematical statistics involving joint densities. Reference 9, pp. 298-309, discusses this type of case. Also, under Assumption 3, the specified index of maintainability can be either the mean or a percentile value, because if  $\sigma^2$  is known, a mean or percentile specification can be translated into a specification on  $\theta$ . If the median  $\bar{M}$  is specified,  $\theta$  is directly determinable from the median independently of  $\sigma^2$ .

In Subsection 7.5.9, the data collected in this study are analyzed and prior distributions for avionic and ground equipments are developed which, in lieu of more applicable procedures, can be used to satisfy Assumption 2. Furthermore, several methods for using predictions, data, and subjective evaluations for establishing a prior distribution for  $\theta$  are discussed. To satisfy Assumption 3, the data and prediction equations presented in Subsection 7.3.5 can be used to provide an estimate of  $\sigma^2$ .

### 7.5.3 Maintainability Index and Test Requirements

The index of maintainability that can be used for the test is the mean, the median, or a percentile value. It is first necessary, however, to translate requirements based on one of the above indices to equivalent requirements on  $\theta$ .

As for conventional tests, two levels of maintainability are to be specified -- a desirable level (to be denoted by the subscript 0) and a minimum acceptable or undesirable level (to be denoted by the subscript 1). Table XLI shows the conversion of mean, median, and percentile specifications to equivalent specifications on  $\theta$ , assuming that  $\sigma^2$  is known.

Given that two values of  $\theta$ ,  $\theta_0$  and  $\theta_1$  ( $\theta_1 > \theta_0$ ) are specified, the requirements on the Bayesian test that are considered here are as follows:

- Requirement I: If  $\theta \leq \theta_0$ , the probability of passing the test is high.
- Requirement II: If an equipment passes the test, the probability that  $\theta$  is greater than  $\theta_1$  is low.

TABLE XLI  
CONVERSION OF MEAN, MEDIAN, AND PERCENTILE SPECIFICATIONS  
TO  $\theta$  SPECIFICATIONS --  $\sigma^2$  KNOWN

Specified Values	Equivalent Specifications
Mean $\mu_0$ $\mu_1$	$\theta_0 = \ln \mu_0 - \sigma^2/2$ $\theta_1 = \ln \mu_1 - \sigma^2/2$
Median $\tilde{M}_0$ $\tilde{M}_1$	$\theta_0 = \ln \tilde{M}_0$ $\theta_1 = \ln \tilde{M}_1$
$P^{\text{th}}$ Percentile $X_p = T_0$ $X_p = T_1$	$\theta_0 = \ln T_0 - Z_p \sigma$ $\theta_1 = \ln T_1 - Z_p \sigma$

These two test requirements can be stated more precisely as follows: Let  $T_n$  be some calculated statistic based on a sample of  $n$  observations of maintenance time. Let  $T^*$  be some preselected critical value for decision such that the equipment passes if  $T_n \leq T^*$  and fails if  $T_n > T^*$ . Then the above two requirements can be written as

$$P[T_n \leq T^* | \theta = \theta_0] = 1 - \alpha \quad (1)$$

$$P[\theta > \theta_1 | T_n \leq T^*] \leq \beta_b \quad (2)$$

If it is assumed that  $\alpha$  and  $\beta_b$  are specified, the only unknowns in the above two equations are  $n$ , the sample size, and  $T^*$  the critical value for the statistic  $T_n$ . The objective of the Bayesian analysis is to determine these two values.

The requirement expressed by Equation 1 is the usual producer's risk control, whereby there is a high probability of acceptance if the mean maintenance time or man-hour value is at the desirable level.

The requirement expressed by Equation 2 offers the consumer assurance that the maintainability level of accepted equipment at least meets a minimum requirement. The notation  $\beta_b$  is used to distinguish the Bayesian risk from the classical Beta risk. In

actuality, Equation 2 represents a control on the upper  $(1-\beta_b)$  percentile of the posterior distribution of the parameter  $\theta$ .

A test based on the requirements I and II is believed to represent the viewpoints of the producer and consumer better than the conventional test. The producer, in his own best interest, will attempt to provide equipment that equals or betters the  $\theta_0$  value, but he would like high assurance that if he does, his equipment will not be rejected. This assurance is provided by requirement I in the same manner as conventional procedures.

The consumer's best interests are served by the more direct approach of assuring acquisition of satisfactory products (in a distributional sense) rather than controlling the probability of accepting poor product. This direct control is provided by requirement II.

#### 7.5.4 General Bayesian Formulation of the Test

Let  $\theta$  be a parameter of a probability density function that has a prior density  $g(\theta)$ . Let  $T_n$  be a statistic based on  $n$  observations whose distribution depends on  $\theta$ . From Bayes' Theorem, the posterior density of  $\theta$  is

$$f(\theta|T_n) = \frac{1}{K} L(T_n|\theta) g(\theta) \quad (3)$$

where

$$K = \int_{-\infty}^{\infty} L(T_n|\theta) g(\theta) d\theta$$

$$L(T_n|\theta) = \text{likelihood of } T_n \text{ given } \theta$$

Equation 3 provides a means for evaluating the distribution of  $\theta$  given the statistic  $T_n$ . In designing an accept/reject demonstration test that is to meet requirements of the form represented by Equation 2, the sample size,  $n$ , and the critical value of  $T_n$ , say  $T^*$ , are to be determined beforehand on the basis of the knowledge that the decision process will be such that  $T_n \leq T^*$  will lead to acceptance and  $T_n > T^*$  will lead to rejection. This type of Bayesian consideration involving future decision actions based on the results of testing is a form of preposterior analysis as defined in Reference 9, page 70.

From this viewpoint, the "given" portion of the posterior density of  $\theta$  can be extended to be the information  $T_n \leq T^*$ , and the posterior density to consider is defined to be

$$f(\theta | T_n \leq T^*) = \frac{1}{K} \int_{-\infty}^{T^*} L(T_n | \theta) g(\theta) dT_n \quad (4)$$

where

$$K = \int_{-\infty}^{\infty} \int_{-\infty}^{T^*} L(T_n | \theta) g(\theta) dT_n d\theta$$

#### 7.5.5 Derivation of the Posterior Density

In this subsection, a closed-form expression for  $f(\theta | T_n \leq T^*)$  is derived under the assumptions listed above. Because of the normality of  $\ln X$ , the natural statistic on a test of  $\theta$ , the expected value of  $\ln X$ , is the arithmetic mean, which is defined by the equation

$$T_n = \frac{\sum_{i=1}^n \ln X_i}{n} = \frac{\sum_{i=1}^n z_i}{n} \quad (5)$$

where  $z_i = \ln X_i$ . Then the likelihood of  $T_n$  is

$$L(T_n | \theta) = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} e^{-\frac{n}{2\sigma^2} (T_n - \theta)^2} \quad (6)$$

From Equation 4, if  $T^*$  is the critical value, and  $y(\theta) \sim N(\bar{\theta}, w^2)$ , then

$$f(\theta | T_n \leq T^*) = \frac{1}{K} \frac{1}{w\sqrt{2\pi}} e^{-\frac{(\theta - \bar{\theta})^2}{2w^2}} \int_{-\infty}^{T^*} \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} e^{-\frac{n}{2\sigma^2} (T_n - \theta)^2} dT_n \quad (7)$$

where

$$K = \int_{-\infty}^{\infty} \frac{1}{w\sqrt{2\pi}} e^{-\frac{(\theta - \bar{\theta})^2}{2w^2}} \int_{-\infty}^{T^*} \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} e^{-\frac{n}{2\sigma^2} (T_n - \theta)^2} dT_n d\theta \quad (8)$$

To evaluate Equations 7 and 8, the following approximation for the cumulative normal probability function, Reference 10, will be used:

$$\int_{-\infty}^v \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} dv = \frac{1}{2} \left[ 1 + \gamma(v) \left( 1 - e^{-\frac{2}{\pi} v^2} \right)^{1/2} \right] \quad (9)$$

where

$$Q(V) = \begin{cases} -1 & \text{if } V \leq 0 \\ 1 & \text{if } V > 0 \end{cases}$$

After the normalizing transformations

$$t_n = \frac{\sqrt{n} (T_n - \theta)}{\sigma} \text{ and } y = \frac{\theta - \theta}{w}$$

are made and the cumulative normal approximation (Equation 9) is applied, Equation 7 reduces to

$$\begin{aligned} f(\theta | T_n \leq T^*) &= f(y | T_n \leq T^*) \\ &= \frac{1}{2K} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \left[ 1 + Q(B-yw) \left( 1 - e^{-A(B-yw)^2} \right)^{1/2} \right] \end{aligned} \quad (10)$$

where

$$A = \frac{2n}{\pi w^2} \text{ and } B = T^* - \theta$$

From the defining normalizing property of K,

$$\begin{aligned} 2K &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \left[ 1 + Q(B-yw) \left( 1 - e^{-A(B-yw)^2} \right)^{1/2} \right] dy \\ &= 1 + \int_{-\infty}^{\infty} Q(B-yw) \frac{e^{-y^2/2}}{\sqrt{2\pi}} \left( 1 - e^{-A(B-yw)^2} \right)^{1/2} dy. \end{aligned}$$

To evaluate K, the series expansion is used:

$$\begin{aligned} (1-x)^{\frac{1}{2}} &= 1 - x/2 - x^2/8 - \dots - c_k x^k - \dots, \quad |x| < 1 \\ &= 1 - \sum_{k=1}^{\infty} c_k x^k \end{aligned}$$

where  $c_1 = 1/2$ ,  $c_k = \left| \frac{(1.5 - k)c_{k-1}}{k} \right|$ ,  $k = 2, 3, \dots$  for approximating the term  $\left( 1 - e^{-A(B-yw)^2} \right)^{1/2}$ .

After squares are completed and the cumulative normal approximation is used, the final result for the posterior density is as follows:

$$f(\theta | T_n \leq T^*) =$$

$$\frac{1}{\sqrt{2\pi}w} e^{-\frac{(\bar{\theta}-\theta)^2}{2w^2}} \left[ 1 + Q(T^* - \theta) \left( 1 - e^{-\frac{2n}{\pi\sigma^2} (T^* - \theta)^2} \right)^{1/2} \right]$$

$$1 + Q(B/w) \left( 1 - e^{-\frac{2}{\pi} \frac{B^2}{w^2}} \right)^{1/2} - \sum_{k=1}^m c_k \sqrt{E_k} e^{-kAB^2 E_k} Q(R_k) \left( 1 - e^{-\frac{2}{\pi E_k} R_k^2} \right)^{1/2}$$

where  $m$  = an integer large enough to yield a good approximation for the series expansion of  $\left( 1 - e^{-A(B-yw)^2} \right)^{1/2}$

$$E_k = (2kAw^2 + 1)^{-1}$$

$$R_k = B/w - 2kABE_k w$$

#### 7.5.6 Derivation of Cumulative Posterior Probabilities

The results of subsection 7.5.5 are now used to develop an expression for the cumulative probability function  $P[\theta \leq \theta_1 | T_n \leq T^*]$ .

Since

$$P[\theta \leq \theta_1 | T_n \leq T^*] = \int_{-\infty}^{\theta_1} f(\theta | T_n \leq T^*) d\theta,$$

from Equation 10

$$P[\theta \leq \theta_1 | T_n \leq T^*] = \frac{1}{2K} \int_{-\infty}^{y_1} \frac{e^{-y^2/2}}{\sqrt{2\pi}} \left[ 1 - Q(B-yw) \left( 1 - e^{-A(B-yw)^2} \right)^{1/2} \right] dy \quad (12)$$

where  $y_1 = (\theta_1 - \bar{\theta})/w$ . To avoid having to use Laplace-Stieltjes integrals because of the nature of the  $Q$  function, the two mutually exclusive cases  $y_1 \leq B/w$  and  $y_1 > B/w$  are considered separately.

Case 1:  $y_1 \leq B/w$

For Case 1,  $Q(B-yw) = 1$  over the range of integration on  $y$  and Equation 12 can be rewritten as

$$P[\theta \leq \theta_1 | T_n \leq T^*] = \frac{1}{2K} \int_{-\infty}^{y_1} \frac{e^{-y^2/2}}{\sqrt{2\pi}} \left[ 1 + (1 - e^{-A(B-yw)^2})^{1/2} \right] dy \quad (13)$$

The result of using the series expansion for the radical and completing the square in the exponent, as was done for evaluating  $K$ , is

$$P[\theta \leq \theta_1 | T_n \leq T^*] =$$

$$\frac{1}{2K} \left[ 2 \int_{-\infty}^{y_1} \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy - \sum_{k=1}^m c_k \sqrt{E_k} e^{-kAB^2 E_k} \int_{-\infty}^{y_1} \frac{1}{\sqrt{2\pi} \sqrt{E_k}} e^{-\frac{1}{2E_k}(y-2kABE_k w)^2} dy \right]$$

Employing Equation 9 after standardizing the normal densities results in

$$P[\theta \leq \theta_1 | T_n \leq T^*] = \frac{1}{4K} \left\{ 2 \left[ 1 - Q(y_1) \left( 1 - e^{-\frac{2}{\pi} y_1^2} \right)^{1/2} \right] - \sum_{k=1}^m c_k \sqrt{E_k} e^{-kAP^2 E_k} (1 + Q(z_k)) \left( 1 - e^{-\frac{2}{\pi z_k^2} z_k^2} \right)^{1/2} \right\} \quad (14)$$

where

$$z_k = y_1 - 2kABE_k w$$

Case 2:  $y_1 > B/w$

For Case 2, the following equation is first considered:

$$P[\theta > \theta_1 | T_n \leq T^*] = \frac{1}{2K} \int_{y_1}^{\infty} \frac{e^{-y^2/2}}{\sqrt{2\pi}} \left[ 1 + Q(B-yw) \left( 1 - e^{-A(B-yw)^2} \right)^{1/2} \right] dy \quad (15)$$

Since  $y_1 > B/w$  over the range of integration,  $Q(B-yw) = 1$ .

Employing exactly the same approach as for Case 1 results in

$$P[\theta > \theta_1 | T_n \leq T^*] = \frac{1}{4K} \sum_{k=1}^m c_k \sqrt{E_k} e^{-kAB^2 E_k} \left[ 1 - Q(Z_k) \left( 1 - e^{-\frac{2}{\pi E_k} Z_k^2} \right)^{1/2} \right] \quad (16)$$

Then

$$P[\theta \leq \theta_1 | T_n \leq T^*] = 1 - P[\theta > \theta_1 | T_n \leq T^*]$$

Substituting for  $K$  produces the following final results:

$$\begin{aligned} & \underline{y_1 \leq B/w} \\ P[\theta \leq \theta_1 | T_n \leq T^*] &= \frac{2 \left[ 1 + Q(y_1) F(y_1) \right] - \sum_{k=1}^m G_k \left[ 1 + Q(Z_k) F\left(\frac{Z_k}{\sqrt{E_k}}\right) \right]}{2 \left[ 1 + Q(B/w) F(B/w) \right] - \sum_{k=1}^m G_k Q(R_k) F\left(\frac{R_k}{\sqrt{E_k}}\right)} \quad (18) \end{aligned}$$

$$\begin{aligned} & \underline{y_1 > B/w} \\ P[\theta \leq \theta_1 | T_n \leq T^*] &= 1 - \frac{\sum_{k=1}^m G_k [1 - Q(Z_k) F(Z_k/\sqrt{E_k})]}{2 \left[ 1 + Q(B/w) F(B/w) \right] - \sum_{k=1}^m G_k Q(R_k) F(R_k/\sqrt{E_k})} \quad (19) \end{aligned}$$

where

$$y_1 = \frac{\theta_1 - \bar{\theta}}{w}$$

$$A = \frac{2n}{\pi \sigma^2}$$

$$B = T^* - \bar{\theta}$$

$$C_1 = 1/2$$

$$C_k = \left| \frac{(1.5-k)C_{k-1}}{k} \right|, k > 1$$

$$E_k = (2kAw^2 + 1)^{-1}$$

$$F(X) = \left( 1 - e^{-\frac{2}{\pi} X^2} \right)^{1/2}$$

$$G_k = c_k \sqrt{E_k} e^{-kAB^2 E_k}$$

$$Q(X) = \begin{cases} -1 & \text{if } X \leq 0 \\ 1 & \text{if } X > 0 \end{cases}$$

$$R_k = B/w - 2kABE_k w$$

$$Z_k = y_1 - 2kABE_k w$$

### 7.5.7 Maintainability-Demonstration-Test Application

The results of Subsection 7.5.6 can now be used to find the sample size  $n$  and critical value  $T^*$  such that

$$P[T_n \leq T^* | \theta = \theta_0] = 1 - \alpha \quad (20)$$

$$P[\theta > \theta_1 | T_n \leq T^*] = \beta_b \quad (21)$$

where

$$T_n = \frac{\sum_{i=1}^n \ln x_i}{n} = \frac{\sum_{i=1}^n z_i}{n}$$

Since  $z_i$  is  $N(\theta, \sigma^2)$ ,  $T_n$  is  $N(\theta, \sigma^2/n)$ , then the requirement expressed by Equation 20 can be rewritten as

$$\int_{-\infty}^{T^*} \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} e^{-\frac{n}{2\sigma^2} (T_n - \theta_0)^2} dT_n = 1 - \alpha \quad (22)$$

or

$$\Phi\left(\frac{T^* - \theta_0}{\sigma/\sqrt{n}}\right) = 1 - \alpha$$

where  $\Phi(\quad)$  denotes the cumulative normal probability function.

If  $Z_\alpha$  equals the normal deviate defined by  $\Phi(Z_\alpha) = 1 - \alpha$ , then, from Equation 22,

$$\frac{T^* - \theta_0}{\sigma/\sqrt{n}} = Z_\alpha$$

or

$$T^* = \theta_0 + Z_\alpha \sigma/\sqrt{n} \quad (23)$$

The second test requirement (Equation 21) is equivalent to the equality

$$1 - P[\theta \leq \theta_1 | T_n \leq T^*] = 1 - \beta_b \quad (24)$$

where  $P[\theta \leq \theta_1 | T_n \leq T^*]$  is obtained from Equation 18 or 19.

Equations 23 and 24 are then sufficient for determining the  $n$  and  $T^*$  values to satisfy the  $\alpha$  and  $\beta$  risks. A generalized computer flow chart for obtaining  $n$  and  $T^*$ , given  $\alpha$ ,  $\beta$ ,  $\theta_0$ ,  $\theta_1$ ,  $\bar{\theta}$ ,  $w$ , and  $\sigma^2$ , is presented in Figure 7.

Standardized sampling-plan tables have been developed by using this flow-chart procedure. In Equation 12, it is seen that  $P[\theta \geq \theta_1 | T \leq T^*]$  depends on only three constants:

$$v_1 = \frac{\theta_1 - \bar{\theta}}{w}$$

$$\frac{B}{w} = \frac{T^* - \bar{\theta}}{w} = \frac{\theta_0 - \bar{\theta}}{w} + \frac{Z_{\alpha}\sigma/\sqrt{n}}{w} \quad (\text{from Equation 23})$$

$$A = \frac{2n}{\pi\sigma^2}$$

For a given  $\alpha$ , the above three constants determine an equivalent set of constants that can be used as indexing parameters:

$$X = \frac{\theta_1 - \theta_0}{n}$$

$$Y = \frac{\bar{\theta} - \theta_1}{w}$$

$$Z = \frac{\theta_1 - \theta_0}{w}$$

Sampling plans in the form  $(n, T^*)$  can then be developed from the entire set  $\alpha$ ,  $\beta$ ,  $Z$ , and  $Y$  to yield a value for  $X$ .

Then, from the definition of  $X$ , the sample size is found from the equation

$$n = \frac{X\sigma^2}{(\theta_1 - \theta_0)^2}$$

and then given  $n$ ,  $T^*$  is found from Equation 23. Table XLII presents the  $X$  values for all combinations of the following parameters:

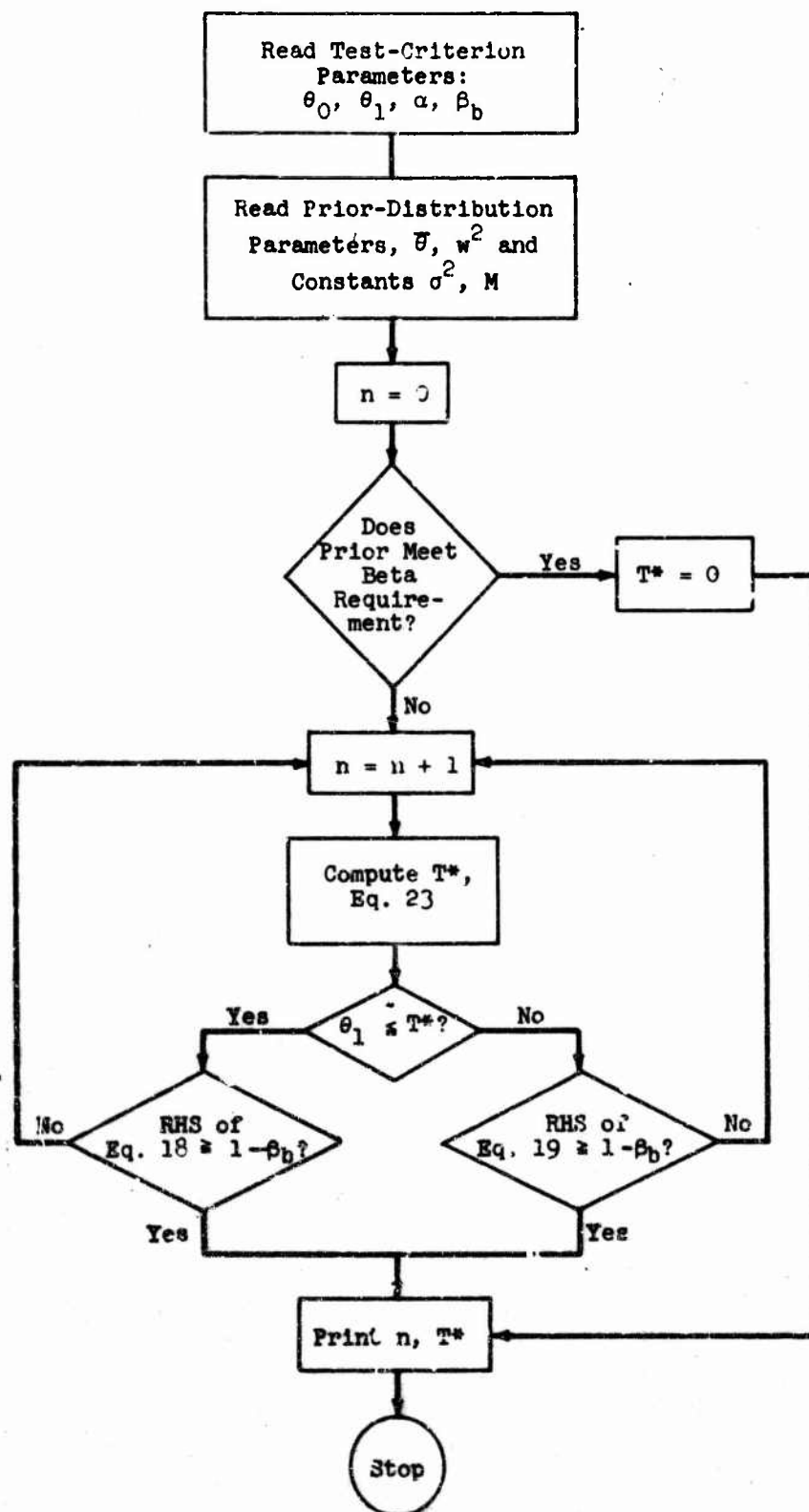


FIGURE 7

FLOW CHART FOR COMPUTER SOLUTION FOR  
SAMPLE SIZE,  $n$ , AND CRITICAL VALUE,  $T^*$

$$\alpha = 0.05, 0.10, 0.20$$

$$\beta_0 = 0.01, 0.05, 0.10, 0.20, 0.50$$

$$Y = -1, -0.5, 0, 0.5, 1, 2$$

$$Z = 0.10, 0.25, 0.50, 0.75, 1.0, 1.5, 2.0, 2.5, 3.0, 4.0$$

A value of  $X=0$  in the table signifies that the prior distribution already meets the Bayesian risk requirement, thus obviating the need for testing.

#### 7.5.8 Example

Assume that corrective-maintenance time for an avionic equipment is lognormally distributed. A maintainability-demonstration test is to be performed with the requirement that if the mean corrective-maintenance time  $\mu$  is 1/2 hour, there will be a 95% acceptance probability. There will also be 99% assurance that an accepted product will not exhibit a mean repair time greater than one hour.

An estimation procedure leads to an estimate of 0.75 for  $\sigma^2$ . Prior information (e.g., that from Table XLIII) indicates that the normal parameter values  $\bar{\theta} = -0.45$  and  $w^2 = 0.30$  can be used for the prior density.

The inputs resulting from the above are as follows:

$$\alpha = 0.10, \beta_0 = 0.01, \mu_0 = 1/2, \mu_1 = 1.0$$

From the equation  $\theta = \ln \mu - \sigma^2/2$ ,

$$\theta_0 = \ln 0.50 - 0.75/2 = -1.068$$

$$\theta_1 = \ln 1.0 - 0.75/2 = -0.375$$

Thus the test requirements are

$$P[T_n \leq T^* | \theta = -1.068] = 0.95$$

$$P[\theta \geq -0.375 | T_n \leq T^*] = 0.01$$

where

$T_n$  = the mean of the natural logarithm of  $n$  observed corrective-maintenance time

The inputs needed to use Table XLII are as follows:

$$\alpha = 0.05$$

$$\beta_b = 0.01$$

$$Y = \frac{\bar{\theta} - \theta_1}{w} = \frac{-0.45 + 0.375}{0.548} = -0.137$$

$$Z = \frac{\theta_1 - \theta_0}{w} = \frac{-0.375 + 1.068}{0.548} = 1.265$$

For conservatism, the next higher tabular entries of  $Y = 0$  and  $Z = 1.5$  were used, leading to the result  $X = 11.275$ . Then

$$n = \frac{X\sigma^2}{(\theta_1 - \theta_0)^2} = \frac{(11.275)(0.75)}{(0.693)^2} = 18$$

The critical value is then

$$\begin{aligned} T^* &= \theta_0 + Z_\alpha \sigma / \sqrt{n} \\ &= -1.068 + 1.645 (0.866/4.242) \\ &= -0.732 \end{aligned}$$

Thus, a random sample of 18 corrective maintenance actions are observed. The sample mean

$$T_n = \frac{\sum_{i=1}^n \ln X_i}{n}$$

is computed. If  $T_n \leq -0.732$ , the equipment is accepted; otherwise a reject decision is made.

TABLE XLII

X VALUES FOR CALCULATING SAMPLE SIZE, n

$$n = \frac{\sigma^2}{(\theta_1 - \theta_0)^2} X$$

Y = -1.000

ALPHA = 0.200

Z	*****		B E T A	*****	
	0.500	0.200	0.100	0.050	0.010
0.10	0.0	0.0	0.020	0.100	0.870
0.25	0.0	0.0	0.060	0.343	1.747
0.50	0.0	0.0	0.160	0.645	2.605
0.75	0.0	0.0	0.255	0.911	3.182
1.00	0.0	0.0	0.350	1.182	3.648
1.50	0.0	0.0	0.515	1.651	4.430
2.00	0.0	0.0	0.633	2.069	5.135
2.50	0.0	0.0	0.706	2.466	5.807
3.00	0.0	0.0	0.774	2.859	6.468
4.00	0.0	0.0	1.367	3.630	7.750

ALPHA = 0.100

Z	*****		B E T A	*****	
	0.500	0.200	0.100	0.050	0.010
0.10	0.0	0.0	0.040	0.190	1.410
0.25	0.0	0.0	0.140	0.600	2.652
0.50	0.0	0.0	0.325	1.124	3.786
0.75	0.0	0.0	0.491	1.485	4.514
1.00	0.0	0.0	0.656	1.802	5.079
1.50	0.0	0.0	0.944	2.438	5.978
2.00	0.0	0.0	1.182	2.957	6.752
2.50	0.0	0.0	1.356	3.425	7.477
3.00	0.0	0.0	1.478	3.868	8.181
4.00	0.0	0.0	1.686	4.724	9.557

ALPHA = 0.050

Z	*****		B E T A	*****	
	0.500	0.200	0.100	0.050	0.010
0.10	0.0	0.0	0.060	0.312	2.063
0.25	0.0	0.0	0.230	0.895	3.568
0.50	0.0	0.0	0.528	1.444	4.974
0.75	0.0	0.0	0.739	2.164	5.846
1.00	0.0	0.0	1.017	2.543	6.495
1.50	0.0	0.0	1.411	3.256	7.907
2.00	0.0	0.0	1.748	3.883	8.348
2.50	0.0	0.0	2.029	4.423	9.116
3.00	0.0	0.0	2.249	4.915	9.870
4.00	0.0	0.0	2.523	5.840	11.281

TABLE XLII (continued)

 $Y = -0.500$ 

ALPHA = 0.200

Z	*****		B E T A		*****	
	0.500	0.200	0.100	0.050	0.010	
0.10	0.0	0.020	0.090	0.290	1.449	
0.25	0.0	0.060	0.324	0.739	2.566	
0.50	0.0	0.170	0.655	1.403	3.623	
0.75	0.0	0.277	0.971	1.911	4.355	
1.00	0.0	0.389	1.290	2.348	4.956	
1.50	0.0	0.566	1.853	3.118	5.994	
2.00	0.0	0.697	2.370	3.821	6.945	
2.50	0.0	0.896	2.868	4.492	7.857	
3.00	0.0	1.234	3.356	5.145	8.738	
4.00	0.0	1.789	4.275	6.358	10.401	

ALPHA = 0.100

Z	*****		B E T A		*****	
	0.500	0.200	0.100	0.050	0.010	
0.10	0.0	0.030	0.170	0.489	2.210	
0.25	0.0	0.120	0.545	1.230	3.720	
0.50	0.0	0.315	1.093	2.107	5.048	
0.75	0.0	0.494	1.499	2.788	5.927	
1.00	0.0	0.676	1.891	3.340	6.625	
1.50	0.0	0.997	2.631	4.254	7.782	
2.00	0.0	1.258	3.256	5.058	8.806	
2.50	0.0	1.443	3.835	5.806	9.773	
3.00	0.0	1.609	4.388	6.523	10.715	
4.00	0.0	2.266	5.450	7.891	12.510	

ALPHA = 0.050

Z	*****		B E T A		*****	
	0.500	0.200	0.100	0.050	0.010	
0.1	0.0	0.030	0.250	0.724	2.985	
0.2	0.0	0.200	0.805	1.751	4.882	
0.50	0.0	0.487	1.564	2.825	6.436	
0.75	0.0	0.755	2.139	3.690	7.465	
1.00	0.0	1.008	2.571	4.353	8.245	
1.50	0.0	1.450	3.432	5.409	9.510	
2.00	0.0	1.826	4.175	6.299	10.600	
2.50	0.0	2.131	4.820	7.112	11.615	
3.00	0.0	2.360	5.437	7.882	12.595	
4.00	0.0	2.664	6.594	9.350	14.482	

TABLE XLII (continued)

 $Y = 0.0$ 

ALPHA = 0.200

Z	*****		B E T A	*****	
	0.500	0.200	0.100	0.050	0.010
0.10	0.0	0.060	0.210	0.507	1.989
0.25	0.0	0.230	0.634	1.242	3.289
0.50	0.0	0.547	1.251	2.136	4.518
0.75	0.0	0.842	1.795	2.825	5.399
1.00	0.0	1.169	2.275	3.422	6.136
1.50	0.0	1.769	3.149	4.481	7.438
2.00	0.0	2.337	3.956	5.457	8.636
2.50	0.0	2.886	4.726	6.379	9.774
3.00	0.0	3.413	5.453	7.243	10.857
4.00	0.0	4.374	6.778	8.833	12.869

ALPHA = 0.100

Z	*****		B E T A	*****	
	0.500	0.200	0.100	0.050	0.010
0.10	0.0	0.110	0.384	0.861	2.958
0.25	0.0	0.412	1.014	1.880	4.636
0.50	0.0	0.889	1.849	3.074	6.129
0.75	0.0	1.307	2.573	3.929	7.148
1.00	0.0	1.669	3.174	4.640	7.977
1.50	0.0	2.447	4.207	5.854	9.401
2.00	0.0	3.138	5.131	6.926	10.680
2.50	0.0	3.779	5.996	7.950	11.896
3.00	0.0	4.396	6.823	8.914	13.058
4.00	0.0	5.545	8.335	10.678	15.216

ALPHA = 0.050

Z	*****		B E T A	*****	
	0.500	0.200	0.100	0.050	0.010
0.10	0.0	0.160	0.533	1.222	3.940
0.25	0.0	0.589	1.440	2.559	5.965
0.50	0.0	1.248	2.503	4.028	7.694
0.75	0.0	1.820	3.366	5.041	8.838
1.00	0.0	2.289	4.091	5.856	9.749
1.50	0.0	3.128	5.276	7.206	11.275
2.00	0.0	3.956	6.300	8.380	12.625
2.50	0.0	4.687	7.244	9.467	13.895
3.00	0.0	5.372	8.143	10.499	15.114
4.00	0.0	6.662	9.806	12.411	17.396

TABLE XLII (continued)

 $Y = 0.500$  $\text{ALPHA} = 0.200$ 

Z	*****		B E T A	*****	
	0.500	0.200	0.100	0.050	0.010
0.10	0.020	0.130	0.386	0.752	2.480
0.25	0.060	0.461	0.979	1.726	3.930
0.50	0.180	0.989	1.856	2.826	5.326
0.75	0.317	1.520	2.573	3.685	6.352
1.00	0.450	2.009	3.217	4.436	7.230
1.50	0.698	2.926	4.392	5.782	8.804
2.00	1.077	3.782	5.467	7.008	10.242
2.50	1.468	4.584	6.461	8.142	11.586
3.00	1.821	5.325	7.381	9.188	12.639
4.00	2.430	6.697	9.091	11.146	15.215

 $\text{ALPHA} = 0.100$ 

Z	*****		B E T A	*****	
	0.500	0.200	0.100	0.050	0.010
0.10	0.030	0.220	0.612	1.231	3.605
0.25	0.110	0.725	1.485	2.536	5.420
0.50	0.296	1.461	2.645	3.932	7.072
0.75	0.500	2.136	3.537	4.957	8.234
1.00	0.715	2.755	4.305	5.825	9.209
1.50	1.116	3.850	5.652	7.334	10.909
2.00	1.442	4.844	6.865	8.689	12.449
2.50	1.783	5.772	7.986	9.942	13.892
3.00	2.344	6.638	9.026	11.103	15.239
4.00	3.219	8.189	10.903	13.214	17.742

 $\text{ALPHA} = 0.050$ 

Z	*****		B E T A	*****	
	0.500	0.200	0.100	0.050	0.010
0.10	0.040	0.343	0.870	1.734	4.728
0.25	0.160	1.017	2.067	3.362	6.880
0.50	0.432	2.060	3.448	5.043	8.755
0.75	0.716	2.746	4.515	6.219	10.041
1.00	0.998	3.513	5.396	7.190	11.096
1.50	1.533	4.781	6.894	8.838	12.901
2.00	2.005	5.893	8.217	10.294	14.518
2.50	2.383	6.922	9.440	11.641	16.034
3.00	2.692	7.885	10.577	12.894	17.461
4.00	3.904	9.613	12.615	15.153	20.093

TABLE XLII (continued)

 $\gamma = 1.000$ 

ALPHA = 0.200

Z	B E T A				
	0.500	0.200	0.100	0.050	0.010
0.10	0.040	0.220	0.533	1.043	2.931
0.25	0.160	0.696	1.363	2.193	4.511
0.50	0.439	1.479	2.443	3.489	6.080
0.75	0.719	2.190	3.342	4.518	7.259
1.00	1.063	2.847	4.148	5.427	8.290
1.50	1.745	4.060	5.503	7.044	10.130
2.00	2.390	5.160	6.902	8.483	11.779
2.50	2.986	6.158	8.076	9.786	13.287
3.00	3.542	7.087	9.171	11.002	14.707
4.00	4.605	8.875	11.282	13.363	17.483

ALPHA = 0.100

Z	B E T A				
	0.500	0.200	0.100	0.050	0.010
0.10	0.060	0.386	0.872	1.600	4.177
0.25	0.250	1.078	1.982	3.141	6.122
0.50	0.661	2.080	3.381	4.733	7.938
0.75	1.070	2.974	4.455	5.932	9.255
1.00	1.449	3.762	5.391	6.958	10.383
1.50	2.288	5.179	7.040	8.756	12.365
2.00	3.101	6.450	8.506	10.354	14.143
2.50	3.845	7.601	9.828	11.794	15.766
3.00	4.525	8.649	11.034	13.116	17.275
4.00	5.742	10.582	13.282	15.598	20.153

ALPHA = 0.050

Z	B E T A				
	0.500	0.200	0.100	0.050	0.010
0.10	0.090	0.523	1.223	2.224	5.419
0.25	0.370	1.481	2.640	4.102	7.684
0.50	0.899	2.689	4.330	5.967	9.719
0.75	1.430	3.771	5.568	7.315	11.159
1.00	1.928	4.687	6.617	8.445	12.365
1.50	2.783	6.275	8.430	10.394	14.460
2.00	3.786	7.687	10.028	12.112	16.336
2.50	4.664	8.965	11.471	13.666	18.055
3.00	5.462	10.126	12.783	15.084	19.645
4.00	6.860	12.205	15.161	17.682	22.613

1J72191

TABLE XLII (concluded)

 $Y = 2.000$  $\text{ALPHA} = 0.200$ 

Z	*****		B E T A		*****	
	0.500	0.200	0.100	0.050	0.010	
0.10	0.110	0.478	0.976	1.692	3.844	
0.25	0.456	1.345	2.211	3.195	5.698	
0.50	1.129	2.576	3.722	4.898	7.622	
0.75	1.809	3.635	4.957	6.245	9.114	
1.00	2.449	4.578	6.036	7.419	10.407	
1.50	3.608	6.216	7.890	9.421	12.638	
2.00	4.654	7.651	9.495	11.157	14.580	
2.50	5.657	9.005	11.009	12.788	16.410	
3.00	6.650	10.341	12.501	14.400	18.222	
4.00	8.604	12.964	15.452	17.603	21.865	

 $\text{ALPHA} = 0.100$ 

Z	*****		B E T A		*****	
	0.500	0.200	0.100	0.050	0.010	
0.10	0.170	0.735	1.472	2.483	5.315	
0.25	0.660	1.901	3.104	4.391	7.513	
0.50	1.518	3.461	4.929	6.393	9.685	
0.75	2.397	4.723	6.361	7.925	11.325	
1.00	3.198	5.830	7.594	9.236	12.731	
1.50	4.608	7.724	9.689	11.468	15.155	
2.00	5.830	9.336	11.466	13.366	17.236	
2.50	6.946	10.801	13.085	15.095	19.153	
3.00	8.023	12.215	14.649	16.771	21.017	
4.00	10.139	14.995	17.736	20.097	24.747	

 $\text{ALPHA} = 0.050$ 

Z	*****		B E T A		*****	
	0.500	0.200	0.100	0.050	0.010	
0.10	0.240	1.019	2.020	3.295	6.757	
0.25	0.878	2.489	4.004	5.580	9.260	
0.50	1.960	4.354	6.121	7.846	11.643	
0.75	2.959	5.797	7.719	9.532	13.404	
1.00	3.929	7.040	9.081	10.962	14.905	
1.50	5.572	9.155	11.382	13.382	17.477	
2.00	6.965	10.929	13.312	15.421	19.684	
2.50	8.199	12.502	15.030	17.243	21.678	
3.00	9.353	13.983	16.658	18.977	23.591	
4.00	11.989	16.580	19.850	22.399	27.400	

1J72151

### 7.5.9 Prior-Distribution Analysis

The prior-distribution input of the parameters  $\bar{\theta}$  and  $w^2$  is a controlling factor in the test. In fact, as shown by the tabular values of Table XLIII,  $X = 0$  if the prior distribution is good. Then, theoretically, no testing is necessary, although in practice this decision would not normally be made.

Prior-distribution development has always been a troublesome area in Bayesian statistics. However, the evidence that Bayesian methods are being used more than ever (e. g., in reliability demonstration) will provide impetus to developing appropriate procedures and data. In addition, the current emphasis on establishing centralized Government and industry reliability and maintainability data banks will provide a good source of historical maintainability data.

At the moment, the following are possible sources of information for use in developing prior distributions:

- . Maintainability predictions and assessments
- . Previous demonstration tests
- . Observed maintainability of similar equipments
- . Design and development tests
- . Subjective evaluations

These sources are discussed briefly in the following paragraphs.

Maintainability predictions and assessments can provide useful data through (1) analysis of past relationships between predicted and observed maintainability, (2) confidence-interval estimates such as those provided through prediction-by-function approaches, and (3) use of several acceptable sets of prediction-input data.

Procedures are available for using the results of previous demonstration tests as prior information for subsequent tests. This is a natural application of Bayesian statistics.

Products similar to those under test are an important data source. Their observed maintainability characteristics are valuable data, but differences in design, part reliability, design maturity, environment, etc., must be accounted for.

As the equipment progresses through development, various engineering tests are performed on parts, assemblies, and components. While these tests may not be designed to provide estimates of maintainability parameters, such estimates can be

obtained if appropriate recording procedures are established. For example, all maintenance times on a design-conformance test should be recorded to establish benchmarks. This information can then be used, for example, to adjust maintainability predictions based on paper designs.

Subjective evaluations are quite controversial, but at this time it is unlikely that sufficient directly applicable data are available for developing prior distributions; engineering judgment will be required to adjust existing data to the problem and to use qualitative appraisals as necessary. In any case, it is important that the producer and consumer agree on the applicability and realism of subjective evaluations used for prior-distribution analysis. Research is being conducted on the quantification of personal judgments for Bayesian applications, and specific approaches have been developed (see Reference 11 for an example).

Three very simple methods for estimating  $\bar{\theta}$  and  $w^2$  are outlined below.

#### 7.5.9.1 Method 1 - Use of Historical Data

The maintenance-time data that were collected for 21 equipments during this study have been analyzed to obtain values for  $\bar{\theta}$  and  $w^2$ , the mean and variance, respectively, for the normal prior density required for use in the Bayesian test. These values are presented in Table XLIII for eight categories representing various combinations of environment, maintenance index, and inclusion or exclusion of no-trouble-found actions.

In the absence of more pertinent data, these values can be used as prior information or can serve as initial values to be modified by predictions, subsequent development tests, and subjective-type analyses.

#### 7.5.9.2 Method 2 - Use of Predictions

Maintainability-prediction procedures such as those presented in MIL-STD-472 and in Volume 1 of this report constitute a means for obtaining a normal prior if "pessimistic" and "optimistic" prediction inputs can be reasonably calculated so that pessimistic and optimistic predictions will be obtained for the mean  $\mu$  or median,  $M$ .

It is recalled that  $\sigma^2$  is assumed to be known either through use of the tabular values given in Table XXVI or through the prediction procedure of Subsection 7.3.5.

TABLE XLIII

OBSERVED MEAN AND VARIANCE OF  $\theta$  FOR  
NORMAL PRIOR-DISTRIBUTION ANALYSIS\*

System Type	No-Trouble-Found Actions Included		No-Trouble-Found Actions Excluded	
	Mean, $\bar{\theta}$	Variance, $w^2$	Mean, $\bar{\theta}$	Variance, $w^2$
Avionic Systems				
Corrective-Maintenance Time	-0.451	0.299	-0.297	0.287
Corrective-Maintenance Man-Hours	0.007	0.550	0.165	0.540
Ground Systems				
Corrective-Maintenance Time	-0.442	0.214	-0.545	0.190
Corrective-Maintenance Man-Hours	0.058	0.226	0.129	0.167
*All $\bar{\theta}$ figures shown are based on times recorded in hours. To convert to minutes, add 4.094 to the tabular value. No conversion is necessary for the variance values.				

Given two predictions of the mean -- say,  $\mu_L$ , the lower or pessimistic value, and  $\mu_U$ , the upper or optimistic value -- there are two equivalent predictions of  $\theta$ :

$$\theta_L = \ln \mu_L - \sigma^2/2$$

$$\theta_U = \ln \mu_U - \sigma^2/2$$

If median values are predicted -- say  $\tilde{M}_L$  and  $\tilde{M}_U$  -- then

$$\theta_L = \ln \hat{M}_L$$

$$\theta_U = \ln \hat{M}_U$$

If it is assumed that the range  $(\theta_U - \theta_L)$  encompasses  $100\alpha(1-p)$  percent of the total of possible values of  $\theta$  and that the best estimate is at the midpoint of the range, the following prior estimates can be used:

$$\bar{\theta} = \frac{\theta_U - \theta_L}{2}$$

$$w^2 = \frac{(\theta_U - \theta_L)^2}{Z_{p/2}^2}$$

where  $Z_{p/2}$  is the normal deviate corresponding to the  $(1-p/2)$  percentile.

Another procedure is to use a predicted value for  $\mu$ , say  $\hat{\mu}$ , and  $100\alpha(1-p)\%$  confidence limits for  $\sigma^2$  from the prediction procedure of Subsection 7.3.5. If  $\sigma_L^2$  and  $\sigma_U^2$  represent the  $100\alpha(1-p)\%$  confidence limits, then

$$\theta_L = \ln \hat{\mu} - \sigma_U^2/2$$

$$\theta_U = \ln \hat{\mu} - \sigma_L^2/2$$

and, as above,

$$\bar{\theta} = \frac{\theta_U - \theta_L}{2}$$

$$w^2 = \frac{(\theta_U - \theta_L)^2}{Z_{p/2}^2}$$

where  $Z_{p/2}$  in this case represents the normal deviate corresponding to a  $100(1-p)\%$  confidence level associated with the  $\sigma_L^2$  and  $\sigma_U^2$  values. The value of  $\sigma^2$  to use in the test should be the point estimate  $\hat{\sigma}^2$  unless there is reason for using a more conservative value.

### 7.5.9.3 Method 3 - Subjective Methods

Instead of predictions, subjective evaluations may be used for obtaining  $\bar{\theta}$  and  $w^2$ . For example, suppose the following is believed to be reasonable:

- (a) There is a 50-50 chance that the mean corrective-maintenance time is less than 1 hour.
- (b) There is only a 10-percent chance that the mean is over 1.5 hours.
- (c) There is only a 1-percent chance that the mean is less than 1/2-hour.

From the relationship  $\mu = e^{\theta + \sigma^2/2}$ , the following are equivalent statements:

- (a)  $P[e^{\theta + \sigma^2/2} \leq 0.5] = 0.01$
- (b)  $P[e^{\theta + \sigma^2/2} \leq 1.0] = 0.50$
- (c)  $P[e^{\theta + \sigma^2/2} \leq 1.5] = 0.90$

Assume a value of 0.8 for  $\sigma^2$ .

Taking logarithms and substituting numerical values leads to the following:

- (a)  $P[\theta \leq -1.093] = 0.01$
- (b)  $P[\theta \leq -0.4] = 0.50$
- (c)  $P[\theta \leq 0.006] = 0.90$

From the normal prior assumption for  $\bar{\theta}$ , the (b) relationship establishes that  $\bar{\theta} = -0.4$ . Two possible  $w^2$  values are as follows:

From (a) and (b)

$$\frac{-1.093 - (-0.4)}{w} = z_{0.99} - z_{0.50} = -2.33$$

or

$$w^2 = 0.088$$

From (b) and (c)

$$\frac{0.006 - (-0.4)}{w} = z_{0.10} - z_{0.50} = 1.28$$

or

$$w^2 = 0.101$$

Averaging these two values yields an estimate of  $w^2 = 0.095$ .

#### 7.5.9.4 Comment

It is emphasized that the three methods described above are quite simplified and that, in fact, a combination of all three may well be used in conjunction with maintenance data that may be available on similar equipment.

If a Bayesian test of the type described is to be performed, management must ensure that necessary tests, data collection, and data analyses take place during the development program for use in establishing a prior distribution or modifying one developed early in the program.

## REFERENCES FOR SECTION VII

1. U. Chand, "On the Derivation and Accuracy of Certain Formulas for Sample Sizes and Operating Characteristics of Nonsequential Sampling Procedures", Journal of Research of the National Bureau of Standards, Vol. 47, No. 6, December 1951.
2. C. Eisenhart, M. Hastay, and W. A. Wallis, Editors, Techniques of Statistical Analysis, Statistical Research Group, Columbia University, McGraw Hill Book Co., 1947.
3. A. Hald, Statistical Theory with Engineering Applications, John Wiley and Sons, Inc. 1952.
4. G. B. Wetherill, Sequential Methods in Statistics, Methuen Monograph, John Wiley and Sons, Inc., 1960.
5. G. Resnikoff and G. Lieberman, Tables of the Non-Central t-Distribution, Stanford University Press, 1957.
6. W. B. Nelson, Producers Risk Tables for the Method 4 Maintainability Demonstration Plan of MIL-STD-471, General Electric Corp., Research and Development Center, Schenectady, N. Y., Report No. 68-C-060, March 1968.
7. L. Miller, "Table of Percentage Points of Kolmogorov Statistics," Journal of the American Statistical Association, Vol. 51, No. 273, March 1956, pp. 111-121.
8. A. Wald, Sequential Analysis, John Wiley and Sons, Inc., 1947.
9. H. Raiffa and R. Schlaifer, Applied Statistical Decision Theory, Harvard University, 1961.
10. J. D. Williams, "An Approximation to the Normal Probability Integral," Annals of Mathematical Statistics, 1946
11. R. Winkler, "The Quantification of Judgment: Some Methodological Suggestions", Journal of the American Statistical Association, Vol. 62, No. 320, December 1967.

## SECTION VIII

### TEST ADMINISTRATION AND IMPLEMENTATION

#### 8.1 GENERAL

Various administrative and procedural aspects of the maintainability-demonstration program are discussed in this section. Many of the guidelines and recommendations are a result of a study of previous demonstration-test plans, results, and critiques.<sup>13</sup>

#### 8.2 TEST SCHEDULING

Table XLIV, which is abstracted from Attachment I of Air Force Regulation 80-14, Test and Evaluation of Systems, Subsystems, and Equipment (R&D), summarizes the three major Air Force R&D test categories.

Ideally, a maintainability demonstration should be scheduled for each test category. Category I tests at the equipment level can provide information for improving the maintainability design before large-scale production build-up. Further design changes, integration problems, and maintenance policies and procedures can be evaluated during a Category II test. A category III test will permit as realistic a measure of operational-system maintainability as possible for final verification. Information is also provided in Category III tests for evaluating the adequacy of the maintainability support program in terms of training, technical manuals, failure-reporting procedures, etc.

By the nature and timing of the tests, a Category I test will probably be based on a fault-inducement sampling procedure. Either fault inducement or natural failures, or both, are applicable for a Category II test, while natural failures should be the primary sampling approach for a Category III test.

In scheduling the maintainability-demonstration test, consideration must be given to other test requirements (e.g., reliability demonstration) and the usually limited number of equipments or systems available for testing in the early stages. The tests must be conducted early enough so that sample-size requirements can be met and time is available for instituting necessary design or procedural changes as a result of the test. On the other hand, the tests should be scheduled so that necessary training has taken place, documentation is complete, and information for establishing fault inducement and other procedures is available. Unfortunately, these two requirements may conflict somewhat and a compromise schedule may have to be adopted.

<sup>13</sup> A particularly valuable reference in this regard is the report, Maintainability-Demonstration Results for 12 Electronic Systems, L. V. Rigby, prepared under Contract AFOT(595)547, Philco Corp., WDL Division, December 1965.

**TABLE XLIV**  
**ACQUISITION TESTING - AFR 80-14**

Category I Tests	Category II Tests	Category III Tests
Test Category Description		
Subsystems Development Test And Evaluation	System Development Test and Evaluation	System Operational Test And Evaluation
Development test and evaluation of the individual components, subsystems and, in certain cases, the complete system under control of AFSC. In addition to qualification, the testing provides for redesign, refinement, and reevaluation as necessary including the practicality of using current standard and commercial items. These tests are conducted predominantly by the contractor, but with the Air Force active participation, evaluation, and control.	Development test and evaluation spanning the integration of subsystems into a complete system in as near an operational configuration as practicable under control of AFSC. Suitable instrumentation will be employed to determine the functional capability of subsystems. Cat. II is an Air Force effort with contractor participation, under Air Force direction and control, and with active operating and supporting command participation. Actual test operation and maintenance should be performed by military personnel who have received formal system training.	Test and evaluation of operational systems by operating command. These tests include all components, support items, personnel skills, technical data, etc., and will be performed under as near operational conditions as practicable. Cat. III testing will be conducted, using a configuration as jointly agreed by the operating command and AFSC/AFLC. Cat. III tests will be conducted per a specific test plan. Cat. III testing terminates when preplanned objectives of acquisition plan have been fulfilled.
Management		
Control: Systems Command Participants: Contractor, Using Commands, Support Commands	Control: Systems Command Participants: Contractor, Using Commands, Support Commands	Control: Using Commands Participants: AFSC, Support Commands, Contractor When Required

### 8.3 PRIOR-INFORMATION REQUIREMENTS

The use of prior information in the design, conduct, and evaluation of a maintainability demonstration has been discussed throughout this report. Several of the more important applications are summarized in Table XLV.

It should be a continuing effort of the maintainability-demonstration management group to ensure that all available pertinent information is properly collected, recorded, and analyzed. Specific efforts to obtain necessary prior information should be planned and instituted as necessary (e.g., special tests or collection of data on similar systems).

### 8.4 MAINTENANCE-PERSONNEL SELECTION

To the extent practicable, the personnel involved in the maintainability demonstration should be representative of those expected during normal operation. The best choice would be those who will actually be assigned to the equipment for then the specific training and experience received during the demonstration program will be of value in the future. Achieving representativeness involves evaluation of skill level, education, general maintenance training and experience, and training and experience that are specific to the equipment in question. These evaluations should apply to both technicians and supervisory personnel.

It is most desirable that the selections be made from Air Force personnel, and this should be stated in the contract unless circumstances prohibit such a clause. If this is made a contractual requirement, the necessary planning for selection, training, and indoctrination can be completed early enough so that even a Category I test can be performed with Air Force personnel. The contractor will normally require that he be allowed approval of selected personnel; therefore, biographical information should be made available to him.

Another consideration concerns the number of personnel to make up the maintenance-team pool. The greater the number, the better the chances for obtaining a representative sampling of maintenance times when technicians are assigned to tasks randomly. In addition to cost factors there is a limit to the number of technicians that should be available. If there are too many technicians, with the limited sample size of the test, each technician will perform only several tasks and thus the learning that would be acquired in operational use is held to a minimum.

It is recommended in the Philec Report<sup>14</sup> that a maintenance team perform no more than six tests and that team personnel not

<sup>14</sup>Ibid. p. 61

TABLE XLV

## SUMMARY OF PRIOR-INFORMATION APPLICATIONS AND DATA SOURCES

Application	Type of Information Required	Sources
Maintainability-Index Selection	Overall system requirements	Concept and Definition Phase studies Technical development and program package plans
Maintainability-Demonstration Requirements	Data for determining realistic and consistent index values	Maintainability predictions Historical data Allocation studies
Maintainability-Demonstration Risk Values	Test and operational cost data	Historical cost data Cost Predictions Logistic analyses
Selection of Statistical Test	Maintainability-distribution data and availability of test inputs	Historical data Engineering analyses Maintainability predictions and analyses
Sample Size and Decision Criterion	Expected values of maintainability-related parameters (e.g., $\sigma^2_{lnAKT}$ )	Maintainability predictions and analyses Historical data
Task-Sampling Scheme	Maintenance-task identification and relative frequency of occurrence	Engineering analyses Maintainability predictions and analyses Historical data
Conduct of Test	Information for achieving maximum possible realism	Operational plans Training-requirement studies Environmental studies

be mixed. The limitation of six tests is, perhaps, somewhat severe. If a maintenance team consists of two men, and 50 tasks are to be performed (e.g., per Test Method 2, MIL-STD-471), there would have to be at least nine teams, or a total of 18 maintenance personnel.

There is no single correct solution. Consideration must be given to the training and equipment-familiarization program, frequency of expected maintenance, planned manning levels, and the learning effects of time compression. From past demonstrations involving approximately 50 tasks each, four or five teams appear to be a reasonable number.

#### 8.5 DEMONSTRATION-REVIEW TEAM

The team designated as responsible for the conduct of the test, and for observation and interpretation of test results, must be carefully selected. Such a team would normally consist of representatives from Air Force Contracts Administration, the equipment contractor, the procuring activity, and the contractor. Representatives of the using command, AFLC, and centralized Air Force maintainability- or effectiveness-assurance offices may also be on the team.

An Air Force representative should be selected as the review-team director; he should be made responsible for test preparation, overall direction of the test, coordination of the review-team activities, and preparation and submission of a report of the demonstration results.

The review team is responsible for the following:

- Observing all aspects associated with the maintenance-task occurrences and determining whether they are valid for demonstration purposes
- Recording maintenance-task-time data and other information pertinent to the decision criterion
- Making decisions to handle unexpected circumstances that may arise
- Obtaining and recording data applicable to satisfying any secondary objectives
- Evaluating the results of the demonstration, preparing the final report, and recommending acceptance or rejection

Members of the review team should be selected with care and should be thoroughly briefed. Particular emphasis must be placed

on the importance of unbiasedness and adherence to the randomness requirement of the sampling process (e.g., in order of task occurrence, technician assignment, etc.).

It is most helpful to have written guidelines for the review team, including the necessary data-collection forms, definitions, and procedures, so that rigor and uniformity are established. Because of a natural variability in observer interpretation and focus, it is generally advisable to have at least two review-team members as task-time-data recorders for each sampled task.

#### 8.6 DEMONSTRATION-DATA FORMS AND RECORDS

There are two basic forms that are pertinent to the demonstration: (1) observer record of individual maintenance-task performance, and (2) summary record of maintenance-task performance.

In addition, there may be a standard form for maintenance-personnel biographies, a checklist for review-team members to verify that all requirements are being met, and other forms to provide supplementary information or meet secondary objectives. As discussed in Section III, the observation of maintenance design or procedural deficiencies is often an important byproduct of the demonstration tests, and plans should be made to identify areas for improvement.

The observer record of maintenance-task performance is the basic data-collection form on which actual times are recorded by the review-team member for each maintenance task. The summary record is the form that summarizes the task-time information from the observer records for use in analyzing the results and calculating the decision statistic of the test.

Because the content and design of such forms depends on the specific conditions of the test, no specific form is recommended. A sample form that can be used as a guide is presented as Exhibit 3. The time data in this sample are recorded by the running-clock method, and the associated actions are described in a narrative manner. After completion of the task, each action can be coded according to preselected categories such as preparation, fault location, repair, and checkout. A debriefing to the maintenance technicians by the review team is recommended, and a special form can be prepared for this purpose.

The summary form contains general background information and the task numbers and desired time elements from the data of the individual task-record forms. An example is shown as Exhibit 4.

An example of a personnel data sheet (taken from the Philco report<sup>15</sup>) used in previous demonstrations of Air Force equipment is presented as Exhibit 5.

Contract No. \_\_\_\_\_

Observer H. Jones

Task No. 4

Test Location Site XYZ

Date 1/1/70

Maintenance Personnel A. Smith

B. Brown

LOCATION OF FAULT

Subsystem \_\_\_\_\_

Equipment \_\_\_\_\_

Assembly \_\_\_\_\_

Part \_\_\_\_\_

FAILURE SIMULATION

Type of Failure \_\_\_\_\_

Method of Simulation \_\_\_\_\_

Observed Symptoms \_\_\_\_\_

Operation Mode \_\_\_\_\_

TIME DATA

Time (Minutes)	Number of Men	Action	Code

Observer Comments \_\_\_\_\_

EXHIBIT 3

RECORD OF MAINTAINABILITY-DEMONSTRATION TASK TIME

Contract Number \_\_\_\_\_ Contractor \_\_\_\_\_

## Numerical Maintainability Demonstration Requirement

**Test Decision Criterion** \_\_\_\_\_

[illegible]

## TASK-TIME SUMMARY RECORD

NAME _____		Age _____	Date _____
Task Rep	Job Classification _____		Type of Work _____
Military Rank _____	Type of Work _____		
Educational Record (How grade completed)			
High School _____	College _____	Grad. School _____	
Technical Schools and/or courses _____			
(Include specific subsystem training)			
			Duration
<b>MILITARY TRAINING</b>			
Branch of Service _____	Rank of School _____	Branch _____	
<b>EXPERIENCE (Civilian and/or Military)</b>			
Company/Branch of Service	Job Title/Rank	Type of Work	Duration
<b>SITE EXPERIENCE</b>		Date Assigned to Station _____	
Subsystem	Date Assigned	Total Experience (Months)	

**EXHIBIT 5**  
**PERSONNEL DATA SHEET**

## **8.7 PREDEMONSTRATION ACTIVITIES**

Prior to the demonstration, the review-team director is responsible for ensuring that all necessary preparations for the demonstration have been completed. The Maintainability-Demonstration Plan (see Section III) should be used as the basic reference.

The major tasks are described below.

### **8.7.1 Site Visit**

The site should be visited to check the existing overall environment and facilities against those planned. Site personnel should be briefed on the character and purpose of the demonstration to ensure understanding and cooperation. The arrangements made for equipment delivery, installation, and checkout should also be made as applicable.

The arrangements made for the necessary maintenance support material, such as test equipment, tools, manuals, and spares, should also be checked against prepared checklists.

Finally, the arrangements made for the demonstration-review team should be checked, including the observation facilities, quarters (for both the maintenance team and review team), instructional guidelines, data forms, and stop watches.

### **8.7.2 Training and Indoctrination**

Maintenance-team and review-team training and indoctrination should be reviewed to ensure that all important details have been covered. A final briefing just prior to the test is also beneficial. It would be expeditious to provide review-team members at this time with a checklist of items for verifying that the test conditions meet the requirements.

The maintenance team should be made completely aware of the purpose of the test and the procedures to be followed. It should be stressed that it is not their individual performance that is being tested but rather that of equipment maintainability design and supporting materials. Before the test begins, the review team should obtain biographies of the maintenance team to aid in task assignment.

### **8.7.3 Preparation of Task-Analysis Package**

When the test sample size has been established and the maintenance-task population identified, it is necessary to determine the actual faults to be induced when sampling is accomplished by fault inducement. (For a natural-failures test, this effort should be directed to providing the information necessary to verify the adequacy of the observed sampling distribution of tasks.)

By use of the procedures suggested in Section VI, a list of maintenance tasks and associated means for their simulation should be prepared. It is recommended that more tasks than required be prepared in order to anticipate unexpected difficulties in the task simulation. The distribution of the "backup" task should also be in accordance with the procedures of Section VI.

It is suggested that trials be performed for the various fault-selection methods, to ensure that the symptom indications will be as planned.

The sequence of tasks should be ordered randomly (e.g., by a "numbers in the hat" method). The maintenance technicians can then be assigned to the tasks in order except when such an approach is contraindicated. Any review-team assignment should also be made at this time.

To acclimate the technicians, instill confidence, and reduce the chances for unrepresentative results due to initial nervousness, it is recommended that the first task for each technician be designated as a trial task. This should be one that is well within the scope of his training and experience and will not be unusually difficult. The results of these trial tasks should not be included in the analysis but should be handled in the same manner as all following tasks. The technicians should not be informed of the nature of these tasks.

The trial tasks also accomplish a secondary objective -- testing the review-team procedures and observer data forms. After a trial task is performed, the review team should meet to identify and correct any deficiencies in the planned procedures.

The result of this effort is a listing of the tasks to be observed, the fault-inducement means for their simulation, the order of their occurrence, and the assigned technicians and review-team members.

## 8.8 CONDUCT OF THE TEST

This subsection is concerned primarily with the conduct of tests based on fault inducement, since the natural-failure sampling approach does not generally require or permit as much careful control.

A "typical" step-by-step procedure for conducting a fault-simulation test is as follows:

- (1) The review-team director or designated representative ensures that the test area is ready for the test and that necessary observer and maintenance personnel are available.

- (2) The checklist of necessary items is read off to ensure that all necessary supporting material is available and the environment and facilities are satisfactory.
- (3) The equipment (including test equipment) is checked to determine if it is in proper working order.
- (4) The first of the preselected malfunctions<sup>16</sup> is inserted into the equipment in accordance with the prescribed procedure. The symptom-display indications may be checked at this time.
- (5) The maintenance technicians are called in and informed of the equipment trouble in a manner previously agreed upon and as similar as possible to that which would occur under normal field operation. This may include a written operator-complaint record or a verbal description by a designated maintenance supervisor.
- (6) Time recording by the observers in accordance with the prescribed task-time record starts as soon as actual maintenance begins.
- (7) Any unusual hindrance or delay, such as unavailability of a required spare part, should be noted by the observers so that final analysis will provide an unbiased measure of active-maintenance time or man-hours.
- (8) When the maintenance task is completed and the repair has been verified by the review team, the area will be cleared and preparation for another task will begin.
- (9) After each task or series of tasks, the maintenance team should be debriefed while the next task is being prepared, to obtain their comments on the maintainability features of the equipment and supporting material.

The following specific precautions should be taken:<sup>17</sup>

- The maintenance technicians must not witness the fault insertion.
- Only prescribed maintenance manuals and handbooks should be made available to the technicians.
- The maintenance technicians should not be permitted to converse with contractor representatives during the demonstration, but normal supervisory personnel should be made available.
- A careful check should be made to ensure that the fault inducement does not provide abnormally informative clues.

<sup>16</sup>As discussed above, it is recommended that for each new technician or team, a trial task be simulated first.

<sup>17</sup>Some of these items, in slightly different form, appeared in a briefing to review-team personnel assigned to a Navy maintainability-demonstration program.

- The maintenance technicians should not know the possible tasks to be simulated, and the contractor should be unaware of the selected task order.
- The demonstration site should provide sufficient visibility for adequate review-team observation.
- The decision to include repair time for any naturally occurring failures or secondary failures should be made beforehand in case such events occur.
- Data records should be reviewed periodically to verify that appropriate detail and accuracy are being recorded.

One specific problem that can occur during the test is extremely long maintenance times. Occasionally a technician may have great difficulty in diagnosing a failure and after a long period of time he has essentially made no headway. To handle this problem, it is advisable to designate a cutoff time--say an estimated 95th percentile value, by which time the test is stopped unless the review team feels that a productive approach is under way.

The analysis of such occurrences poses problems since the long maintenance times may well be due to inexperience of the technicians. If it is clear that with experience the technician would not have had the difficulty (and perhaps this can be checked by repeating the test using another technician), an estimated value of the expected task time under more realistic conditions may be used. As an alternative, another back-up task may be observed to replace the troublesome one. While these are practical solutions, their implementation opens up the question of test bias.

In case a review of the problem indicates that the fault is with the maintainability design rather than with the technician or environment, an estimate of the required test statistic such as the mean can be made by analysis of the completed-task times in conjunction with statistical theory on censored or truncated observations.

As a simple example, assume that the guidelines to the review team are to stop the test when the maintenance time exceeds  $X^*$  (say an estimated 99th percentile value). If  $n$  observations are made, and  $k$  observations are greater than  $X^*$  ( $k$  should, of course, be small, say only 1 or 2), then actual times are recorded on  $(n-k)$  tasks.

If a lognormal distribution is assumed, the maximum likelihood estimation for the mean is

$$\hat{\mu} = \bar{\ln X} - \frac{1}{n-k} (\bar{\ln X} - \ln X^*)$$

where

$$\overline{\ln X} = \frac{\sum_{i=1}^{n-k} \ln X_i}{n-k}$$

and

$\hat{\gamma}$  is an auxiliary estimating function for which tables are available.<sup>18</sup>

### 8.9 DATA REDUCTION AND ANALYSIS

The primary effort (data reduction and analysis) is the development of the statistics required by the statistical demonstration-test procedures, such as mean number of man-hours to complete corrective maintenance or the time associated with the 95th percentile. This effort should be under Air Force control, preferably through the review-team director.

Before such statistics can be obtained, however, data should be edited to ensure recording accuracy and consistency and to check that all test-condition requirements have been met.

The assumptions made for the statistical-test design can first be checked, and if there is indication of a disparity, further analysis will be required before a decision can be made to proceed. An alternate nonparametric test can be used, for example, if it is found that the lognormal-distribution assumption is a poor one.

In addition to the calculation of the test statistic, further data are generally available that can add to the store of knowledge of maintainability. At a minimum, these data should be tabulated and summarized.

The means for calculating the test statistic and the application of the decision criterion has, of course, been determined beforehand, and once numerical values are obtained, the decision of accepting or rejecting is fairly straightforward.

### 8.10 TEST REPORTING

The final effort of the demonstration activity is the preparation of the Maintainability Demonstration Report by the review team. For extended test periods interim reports may have also been prepared.

<sup>18</sup>A. Clifford Cohen, Jr., "Tables for Maximum Likelihood Estimates: Singly Truncated and Singly Censored Samples", Technometrics, Vol. 3, No. 4, November 1961.

The final report need not be unnecessarily detailed, since reference can be made to the Maintainability Demonstration Test Plan. The following is a possible format for the report:

1. Introductory Section

A summary of test objectives, including identification of equipment, manufacturer, contract number, numerical requirements, demonstration site, and review-team members.

2. Test Conditions

A summary of the test conditions, including maintenance personnel, with particular reference to deviations from the Maintainability Demonstration Plan.

3. Test Procedures

A brief review of the proceedings of the demonstration test, noting particular problems and means taken to overcome them.

4. Test Results

A summary of the observed data and the results of the analysis made for decision purposes.

5. Discussion

A discussion of the test results and analysis, along with qualitative findings of the review team and maintenance team. Deficiencies in test design and procedures should be noted here, as well as deficiencies in the maintainability design and procedures associated with the equipment under test.

6. Recommendations

A specific recommendation on acceptance or rejection of the equipment under test and other recommendations for improvement in equipment, procedures, or test design.

8.11 ACTIONS FOLLOWING REJECTION.

It is a requirement of MIL-STD-471 that a retest be performed if the equipment fails to pass the demonstration test. Such a retest must be scheduled to permit design and procedural changes to be made to correct existing deficiencies. Generally, the cost of such redesign should be borne by the contractor and not be a subject of renegotiation.<sup>19</sup> If this is not stated clearly and unequivocally in the initial contract, one of the major purposes of demonstration -- providing the incentive for good design -- is thwarted.

<sup>19</sup>This, of course, is based on the assumption that test failure is due to contractor deficiencies rather than Government deficiencies.

A more direct means for providing such incentive is through an incentive contract in which the amount of the contract award depends on the results of the demonstration effort.

For the retest, the contractor should first submit a report describing the efforts made to improve the equipment's maintainability, and these efforts should be approved by the procuring activity. A second test of unimproved equipment or procedures is to be avoided.

The task-sampling scheme should be revised for the retest. The same procedure may be applicable, but a new set of random numbers should be selected as applicable so that the contractor cannot, by knowing the tasks to be sampled, plan his improvements solely to pass the test.

All other conditions concerning the conduct of the test should apply for the retest.

## SECTION IX

### CONCLUSIONS AND RECOMMENDATIONS

#### 9.1 GENERAL

This concluding section is, in essence, a summary of the major points made in the report. The major conclusions and recommendations made in the previous sections are repeated or summarized here for emphasis. For convenience, the pertinent sections of the report that deal with the subject of interest are indicated.

#### 9.2 THE CONCEPT OF MAINTAINABILITY DEMONSTRATION

Maintainability demonstration must be a contractual requirement in accordance with MIL-STD-470, and plans for meeting the requirements of MIL-STD-471 must be detailed in a Maintainability Demonstration Plan submitted by the contractor (Section 1.1).

The demonstration procedure is essentially an application of the statistical theory of hypothesis testing and must be planned and conducted as such (Section 1.1).

#### 9.3 THE CURRENT STATUS OF MAINTAINABILITY DEMONSTRATION

Some of the more important conclusions and recommendations resulting from the survey (Section II) are as follows:

- Maintainability demonstration is primarily a maintainability-control and information-generating process.
- Many contractual documents contain inadequate maintainability-demonstration provisions.
- A major problem concerning the conduct of the test is the differences between test and field environments.
- Allocation from higher-level requirements is the preferred method for determining numerical test requirements.
- The mean is the generally preferred index of specification.
- MIL-STD-471 is generally acceptable except that greater flexibility should be offered.

#### 9.4 DEMONSTRATION AS A PROGRAM ELEMENT

Timely planning and careful management of the maintainability-demonstration program are required to fulfill the requirements of MIL-STD-470 and MIL-STD-471 (Section 3.1).

The following additional conclusions are made:

- A maintainability-demonstration test does not guarantee achieving the required maintainability. It focuses the contractor's attention on maintainability, but often this is not sufficient unless penalties for test failure are included in the contract (Section 3.2).
- The contractor's maintainability-demonstration plan must meet the requirements of MIL-STD-471. Its development should be one of continuous refinement to reflect changes in requirements and design and to incorporate the results of maintainability-design reviews, predictions, and assessments (Section 3.3).
- An important by-product of a maintainability-demonstration test is the information provided for improving equipment reliability and maintainability. The passing of the test does not mean that important improvements cannot be made. The procuring activity should plan to have appropriate personnel monitoring the test to discover design deficiencies and recommend improvement (Section 3.4).

#### 9.5 THE STATISTICAL BASIS OF DEMONSTRATION

- Full understanding of the meaning of the  $\alpha$  and  $\beta$  risks associated with a demonstration-test specification must precede the assignment of numerical values (Section 4.2).
- The numerical maintainability-demonstration test requirements must be presented in a manner not subject to misinterpretation. This is especially important when the requirement is stated in terms of confidence levels (Section 4.3).
- The choice between variables and attributes tests, single, multiple, and sequential tests, nonparametric and parametric tests, and classical and Bayesian tests requires full consideration of the information requirements and necessary assumptions associated with combinations selected as possible alternatives (Section 4.6).

#### 9.6 MAINTAINABILITY-DEMONSTRATION-TEST SPECIFICATION

- The more important requisites for a maintainability-demonstration-test specification are as follows (Section 5.1):
  - The maintainability index should represent a measure that is directly influenced by equipment design so that the producer can plan for high assurances of a pass decision, but bears the responsibility for a reject decision.

- Relationships (at least qualitative) between design parameters and the maintainability index should be known so that design evaluations and predictions are possible.
  - The maintainability index should be appropriate for, and measurable in, the demonstration-test environment.
  - The maintainability index should be related to higher-level system-requirement parameters, and numerical values should be consistent with values for these higher-level parameters.
  - Adequate sampling and statistical-evaluation procedures should be available for demonstrating conformance to the requirement.
- Corrective maintenance is generally more critical than preventive maintenance when operational requirements are considered, especially if the latter can be scheduled during non-use periods (Section 5.2).
  - The man-hour-rate index, such as maintenance man-hours per operating hour, is a direct function of both reliability and maintainability; therefore, a test based on such an index should be the responsibility of both the reliability and maintainability groups. Models for relating expected man-hours per maintenance action to man-hour rate can be developed so that a test based on the former provides an indirect control on the latter (Section 5.2).
  - The mean index is strongly influenced by long maintenance times, while the median is not. The mean generally provides better manpower cost control, is derivable from higher-level specifications, and has more desirable statistical properties than the median. The median is applicable to distribution-free tests, has direct operational meaning in terms of being a 50-percent percentile value, and for the lognormal distribution is not dependent on the value of  $\sigma^2$  (Section 5.2).
  - The Maintainability Index Selection Matrix (Exhibit 2) should be used as a guide in choosing the maintainability index (Section 5.2).
  - Three basic criteria for assigning numerical values for the selected maintainability index are (Section 5.3):
    - (1) The specified value should be consistent with higher system-level requirements.

- (2) It should be realistic.
- (3) It should pertain to the demonstration environment.
- Trade-off approaches between reliability and maintainability indices, given an overall availability requirement, should be considered for obtaining numerical requirements (Section 5.3).
- Maintainability predictions and analysis of historical data should be used to assess the realism of specified values. Tables XII, XIII, XIV, and XV present pertinent historical data (Section 5.3).
- The closer the test environment to the expected field environment, the more meaningful the demonstration. Every effort should be made to achieve such similarity. Tables XVI and XVII present information factors to consider and causes of discrepancy, respectively (Section 5.3).
- The assignment of test risks must consider the corresponding effect on required test sample size (Section 5.4).
- Use of prior information in terms of test costs and wrong-decision costs can be used in a decision-theory model for assigning appropriate test risks (Section 5.4).

## 9.7 SELECTION OF MAINTENANCE-TASK SAMPLE

- The natural-failure approach to sample selection is the preferred choice but often is impractical because of time and equipment limitations (Section 6.2).
- A combination of natural-failure and fault-inducement procedures should be considered if it can be feasibly implemented (Section 6.2).
- Failures that occur naturally during a simulated failure test should be included in the sample (Section 6.2).
- Care must be taken to see that the fault-inducement procedure generates a representative sampling of maintenance tasks. Intermittencies, degradation-type failures, secondary failures, and the like should be considered along with the usual catastrophic-failure modes (Section 6.3).
- The contractor should make plans for retaining parts, circuit cards, assemblies, etc., that have been rejected during development, reliability, and quality-control tests for use in inducing noncatastrophic failures (Section 6.3).

- A list of maintenance tasks is necessary for implementing natural-failure tests as well as fault-inducement tests (Section 6.3).
- Simple random sampling is preferred over proportional stratified random sampling when the analytical aspects of the demonstration procedure are considered. Proportional stratified sampling, however, can provide better assurance of a representative sample (See Table XX, Section 6.4).
- The basic criterion for stratifying the population of maintenance tasks is the expected maintenance time (or man-hours); that is, tasks within a stratum should require essentially the same maintenance effort so that variability within a stratum is small (Section 6.4).
- The frequency with which tasks are selected within a stratum is a function of the relative frequency of occurrence of the tasks comprising the stratum (Section 6.4).
- A stratum comprising tasks not associated with piece-part failures and not necessarily a result of unreliability should be developed if the occurrence probability of such tasks is not negligible (Section 6.4).
- The symptom matrix can be used as an approach to stratification, especially if fault-location time is the major element of the maintenance-task times (Section 6.4).

## 9.8 STATISTICAL MAINTAINABILITY-DEMONSTRATION-TEST PLANS

- A review of MIL-STD-471 plans is presented in Subsection 7.2; it should be consulted to determine the applicability of these plans to the specific problem at hand.
- Data and methods for estimating  $\sigma^2$  in ART, the variance of the logarithm of corrective-maintenance time, are presented in Subsection 7.3.5 for use in determining numerical values and sample-size requirements and for use in Bayesian tests.
- Fourteen non-Bayesian maintainability-demonstration plans are presented in Section 7.3. Alternatives with respect to fixed vs. sequential tests, lognormal vs. nonparametric tests, and various parameter specifications are considered. Guidelines for test selection are also presented. These plans are believed to represent improvements over the MIL-STD-471 plans in terms of greater flexibility in risk assignment, test-parameter specification, hypotheses, and forms of testing.

- A Bayesian test is developed and illustrated in Section 7.5 for a fixed-sample test based on a lognormal assumption for the distribution of maintenance time. Sampling-plan tables and methods and data for prior-distribution analyses are also presented.

## 9.9 TEST ADMINISTRATION AND IMPLEMENTATION

- The scheduling of maintainability-demonstration tests should conform to the test categories described in AFR 80-14 to the extent possible (Section 8.2).
- Management should make a continuing effort to acquire and evaluate all available pertinent information applicable to the demonstration (See Table XLV, Section 8.3).
- Air Force personnel are the preferred choice for maintenance-team selection. Representativeness with respect to education, training, skills, and experience must be considered in choosing individuals (Section 8.4).
- A demonstration-review team under the control of an Air Force representative should be established (Section 8.5).
- Example data forms, presented in Section 8.6, can be used as a guide for developing forms applicable to the problem at hand.
- Predemonstration activities by the review team should include a site visit, training and indoctrination review, and preparation of task sample package (Section 8.7).
- For each technician, a trial task should be established for acclimating the technician and instilling confidence (Section 8.7).
- A "typical" step-by-step procedure for conducting the test and a list of precautions is presented in Section 8.8.
- A maintainability-demonstration report should be prepared by the review team. A sample outline is presented in Section 8.10.
- If the equipment fails the test, a retest must be scheduled in accordance with the requirements of MIL-STD-471. The task sampling for the retest should be based on a new set of random numbers (Section 8.11).

# APPENDIX A

## THE LOGNORMAL DISTRIBUTION\*

### 1. BASIC PROPERTIES

A random variable  $X$  has a lognormal probability distribution if the logarithm of the variable is distributed normally. The lognormal probability density function is

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-(1/2\sigma^2)(\ln x - \theta)^2}, \quad 0 < x < \infty$$

where  $\ln X$  = natural logarithm of  $X$ .

If  $Y = \ln X$ , the probability density of  $Y$  is normal, mean  $\theta$ , variance  $\sigma^2$ .

The important moment and distributional properties of the lognormal are presented below:

$$\text{Mean} = e^{\theta + \sigma^2/2}$$

$$\text{Variance} = e^{2\theta + \sigma^2} (e^{\sigma^2} - 1)$$

$$\text{Median } X = \text{Geometric Mean} = e^{\theta}$$

$$\text{Mode } X = e^{\theta - \sigma^2}$$

$$p^{\text{th}} \text{ Percentile} = e^{\theta + Z_p \sigma} \quad (Z_p = \text{normal deviate, i.e.,}$$

$$\int_{-\infty}^{Z_p} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 1-p)$$

$$\text{Coefficient of Variation} = e^{\sigma^2} - 1$$

\*The Lognormal Distribution, by J. Aitchison and J. Brown, Cambridge University Press, 1963, is the most complete reference on this subject.

The lognormal is a positively skewed distribution, with degree of skewness increasing as  $\sigma^2$  increases. From the above formulas, it is seen that mode < median < mean. Because of the two parameters  $\theta$  and  $\sigma^2$ , the lognormal is a relatively flexible distribution. Several important properties possessed by the distribution are given below. The notation  $L_X(\theta, \sigma^2)$  shall be used to represent a lognormal density with parameters  $\theta$  and  $\sigma^2$ .

(A) If  $X = L_X(\theta, \sigma^2)$ , then  $Z = cX^b = L_Z(\ln c + b\theta, b^2\sigma^2)$ , ( $c > 0$ )

(B) If  $X = L_X(\theta, \sigma^2)$ , then  $Z = 1/X = L_Z(-\theta, \sigma^2)$   
[case (A) with  $c=1$ ,  $b=-1$ ]

(C) If  $X_1 = L_{X_1}(\theta_1, \sigma_1^2)$  and  $X_2 = L_{X_2}(\theta_2, \sigma_2^2)$  are independent,

$$Z_1 = a_1 X_1 \cdot a_2 X_2 = L_{Z_1}(a_1 \theta_1 + a_2 \theta_2, a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2), (a_1 > 0, a_2 > 0)$$

$$Z_2 = a_1 X_1 / a_2 X_2 = L_{Z_2}(a_1 \theta_1 - a_2 \theta_2, a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2), (a_1 > 0, a_2 > 0)$$

Properties (A) and (B) state that the lognormality of a random variable is preserved under a simple shape or scale transformation, and property (C) states that the product and quotient of independent lognormal variates is also lognormal.

It can also be shown that if the product of two independent random variables is lognormal, then each random variable must also be lognormal. This is not necessarily true for dependent variates.

The following central limit theorems hold:

(a) If  $X_1, X_2, \dots, X_n$  are independent positive variates with the same distribution and

$$E[\ln X_j] = \theta < \infty \quad j=1, 2, \dots, n$$

$$\text{Var}[\ln X_j] = \sigma^2 < \infty \quad j=1, 2, \dots, n$$

$$Z = \prod_{j=1}^n X_j \text{ is asymptotically } L_Z(\ln n, n\sigma^2)$$

As a corollary, the geometric mean  $(\prod_{j=1}^n X_j)^{1/n}$  is asymptotically lognormal, mean  $= \theta$  variance  $= \sigma^2/n$ .

(B) If  $X_1, X_2, \dots, X_n$  are independent positive variates such that

$$E(\ln X_j) = \theta_j < \infty$$

$$\text{Var}(\ln X_j) = \sigma_j^2 < \infty$$

$$E(|\ln X_j - \theta_j|^2) = \omega_j^2 < \infty$$

then  $Z = \prod_{j=1}^n X_j$  is asymptotically  $L_2$   $\left( \sum_{j=1}^n \theta_j, \sum_{j=1}^n \sigma_j^2 \right)$

provided that  $\left[ \frac{3 \sqrt{\sum_{j=1}^n \omega_j^2}}{\sqrt{\sum_{j=1}^n \sigma_j^2}} \right] \rightarrow 0$  as  $n \rightarrow \infty$

## 2. DATA ANALYSES

### 2.1 Graphical Analysis

It is usually advisable when analyzing any set of data to plot the observed sample observations so that distributional assumptions may be verified.

There exists logarithmic probability paper which yields a straight line for the lognormal cumulative distribution. Thus, if the cumulative distribution of an observed sample plots approximately as a straight line, the lognormal assumption may be accepted as being reasonable. Naturally, such a test is not rigorous but affords a quick method of judging. Such tests as the Chi Square test and the Kolmogorov Smirnov test are preferred when more rigor is required.

The use of lognormal probability paper also affords a quick estimate of  $\theta$  and  $\sigma^2$ . From the quantile equation  $X_p = e^{\theta + Z_p \sigma}$ , we have

$$X_{0.16} = e^{\theta - \sigma}$$

$$X_{0.50} = e^{\theta}$$

$$X_{0.84} = e^{\theta + \sigma}$$

Thus

$$\theta = \ln X_{0.50}$$

and

$$\sigma = \ln \frac{1}{2} \left[ \frac{X_{0.50}}{X_{0.16}} + \frac{X_{0.84}}{X_{0.50}} \right]$$

To obtain estimates of  $X_{0.16}$ ,  $X_{0.50}$  and  $X_{0.84}$ , a straight line is fitted through the points plotted on lognormal probability paper and the  $X$  values corresponding to the cumulative values of 16 percent, 50 percent, and 84 percent are read off. As an example, a simulated sample of 50 from a lognormal distribution with  $\theta = 3$ ,  $\sigma^2 = 0.5$  resulted in the observed cumulative distribution plotted in Figure A-1 on lognormal probability paper.

A straight line fitted through the points yields the estimates  $X_{0.34} = 32.0$ ,  $X_{0.50} = 20.5$  and  $X_{0.16} = 13.0$ . Using the above equations we have the estimates  $\theta = 3.02$  and  $\sigma^2 = 0.45$  corresponding to the true values of 3 and 0.5.

## 2.2 Point Estimates

For the lognormal distribution, it is advisable to consider separately the estimation of  $\theta$  and  $\sigma^2$  from that of the mean  $E(X)$  and the variance  $V(X)$ . Several estimation procedures are possible for each, such as maximum likelihood, method of moments, method of quantiles, and graphical approaches.

Table A-1 presents these various methods of estimation assuming that a random sample of  $n$  observations is obtained. The corresponding variance of the estimates and appropriate comments are also given in the table.

To illustrate the characteristics of the possible estimation procedures, 25 samples of 50 observations each were obtained by computer simulation from a lognormal distribution with  $\theta = 3$  and  $\sigma^2 = 0.25$ . Estimates of  $\theta$ ,  $\sigma^2$ ,  $E(X)$  and  $V(X)$  were computed by the maximum likelihood, moment, and quantile methods for each sample. The average of these sample values as well as the sample standard deviation of these estimates are presented in Table A-2. It is seen that in all cases the estimates are reasonably close to the true value but that the maximum-likelihood estimates have generally the greatest precision in terms of minimum standard deviation.

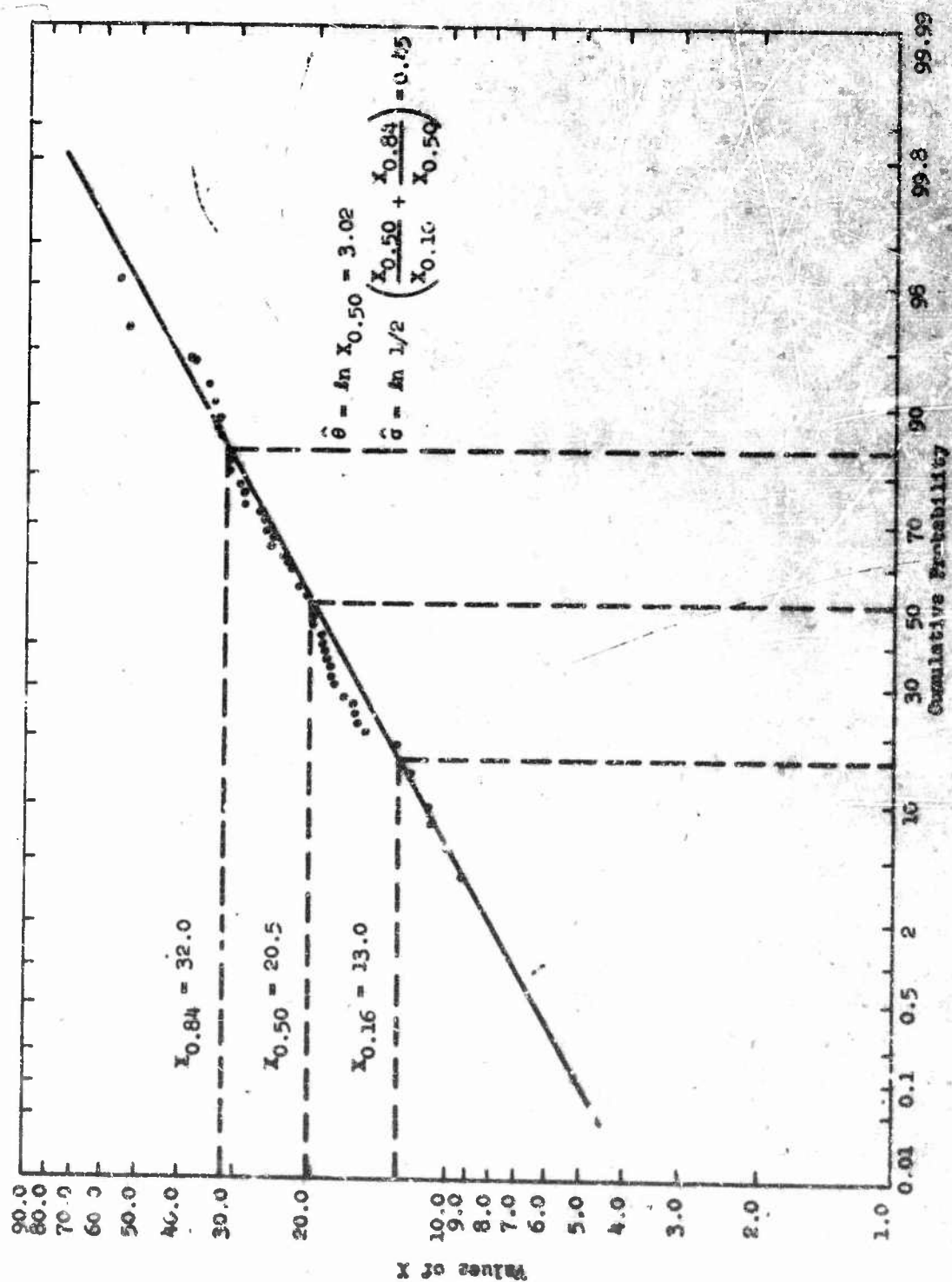


FIGURE A-2  
OBSERVED CUMULATIVE DISTRIBUTION: SAMPLE OF 50 OBSERVATIONS  
SIMULATED FROM  $L_2(3,0.5)$

TABLE A-1  
BEST ESTIMATES AND SOME PROPERTIES FOR PARAMETERS OF THE LOGNORMAL DISTRIBUTION

Parameter	Method	Estimate, $\hat{\theta}$	Variance of Estimate, $V(\hat{\theta})$	Comment on Estimate
$\sigma$	Maximum Likelihood	$\hat{\sigma}_1 = \frac{1}{n} \sum_{i=1}^n \ln X_i$	$\sigma^2/n$	Preferred method-minimum variance.
	Moment	$\hat{\sigma}_2 = 2\ln X - \frac{1}{2} \ln \left( \frac{\sum_{i=1}^n X_i^2}{n} \right)$	$\left( \frac{1}{4n} \right) (\gamma^4 + 4\gamma^3 - 2\gamma^2 + 4\gamma)$	Inefficient compared to M.L.E.
	Quantile	$\hat{\sigma}_3 = \frac{1}{2} (\ln X_{0.27} + \ln X_{0.73})$	$1.23 (\sigma^2/n)$	Acceptable for large n. Easy computation.
	Graphical	$\hat{\sigma}_4 = \frac{1}{2} \ln 2$	Not Applicable	For early indication. Subjective.
$\sigma^2$	Maximum Likelihood	$\hat{\sigma}_1^2 = \frac{1}{n-1} \sum_{i=1}^n (\ln X_i - \hat{\sigma}_1)^2$	$2\sigma^4/(n-1)$	Preferred Method.
	Moment	$\hat{\sigma}_2^2 = \frac{1}{n} \sum_{i=1}^n \ln^2 X_i - 2\ln X$	$\frac{1}{n} (\gamma^4 + 4\gamma^3 + 2\gamma^2)$	Very inefficient compared to M.L.E.
	Quantile	$\hat{\sigma}_3^2 = 0.338 (\ln X_{0.85} - \ln X_{0.07})$	$3.08 (\sigma^4/n)$	Acceptable for large n.
	Graphical	$\hat{\sigma}_4^2 = \left\{ \ln \left[ \frac{1}{2} \left( \frac{X_{0.50}}{X_{0.16}} + \frac{X_{0.84}}{X_{0.50}} \right) \right] \right\}^2$	Not Applicable	For early indication.
Mean, $\bar{X}$	Maximum Likelihood	$\hat{\bar{X}}_1(X) = e^{\hat{\sigma}_1} \bar{X}_n \left( \frac{\hat{\sigma}_1^2}{2} \right)$	$\frac{\sigma^2}{2n} (\gamma^2 + \gamma^2)$	Preferred method. Most computation.
	Moment	$\hat{\bar{X}}_2(X) = \bar{X}$	$\gamma^2 \sigma^2/n$	Acceptable especially for combining samples.
	Quantile	$\hat{\bar{X}}_3(X) = e^{\hat{\sigma}_3} + \hat{\sigma}_3^2/2$	$\left[ 4\gamma^2 v(\hat{\sigma}_3) + \frac{1}{2} v(\hat{\sigma}_3^2) \right] \sigma^2$	Acceptable, minimum efficiency - 65 percent as $\sigma^2 \rightarrow \infty$ .
	Graphical	$\hat{\bar{X}}_4(X) = e^{\hat{\sigma}_4} + \hat{\sigma}_4^2/2$	Not Applicable	Generally good for large n and small $\sigma^2$ .
Variance, $V(X)$	Maximum Likelihood	$\hat{V}_1(X) = e^{2\hat{\sigma}_1} \left[ \bar{X}_n \left( \frac{2\hat{\sigma}_1^2}{n-1} \right) - \bar{X}_n \left( \frac{2\hat{\sigma}_1^2}{n-1} \right) \right]$	$\left[ 4\sigma^2 \gamma^4 + 2\sigma^4 (\gamma^2 + 1)^2 \right] \sigma^4$	Preferred method.
	Moment	$\hat{V}_2(X) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$	$(\gamma^{12} + 6\gamma^{10} + 15\gamma^8 + 16\gamma^6 + 2\gamma^4) \sigma^4$	Very inefficient-should not be used.
	Quantile	$\hat{V}_3(X) = e^{2\hat{\sigma}_3} + \hat{\sigma}_3^2 (e^{\hat{\sigma}_3^2} - 1)$	$\left[ 4\gamma^4 v(\hat{\sigma}_3) + (\gamma^2 - 1)^2 v(\hat{\sigma}_3^2) \right] \sigma^4$	Acceptable-efficiency between 65 percent and 70 percent.
	Graphical	$\hat{V}_4(X) = e^{2\hat{\sigma}_4} + \hat{\sigma}_4^2 (e^{\hat{\sigma}_4^2} - 1)$	Not Applicable	Generally good for large n and small $\sigma^2$ .

NOTES: (1)  $X_p$  = Observed  $p^{th}$  percentile value. Interpolation may be necessary.

(2)  $\hat{\sigma}_p$  =  $p^{th}$  percentile value obtained from straight line fitted through observed points on log-probability paper.

(3)  $\bar{X}_n(t) = \bar{X} \left[ 1 - \frac{1}{2} \left( \frac{t^2}{\sigma^2} \right) + \frac{1}{24} \left( \frac{t^4}{\sigma^4} \right) - \frac{1}{720} \left( \frac{t^6}{\sigma^6} \right) \right]$

(4)  $\gamma^2 = e^{\sigma^2} - 1$

(5)  $\bar{X}^2 = \left[ \bar{X}(X) \right]^2 = e^{2\sigma_1} + \sigma^2$

TABLE A-2

AVERAGE SAMPLE ESTIMATE (E) AND SAMPLE STANDARD DEVIATION  
OF ESTIMATES (S) FOR THE FOLLOWING METHODS

Parameter	Theoretical Value	Symbol	Maximum Likelihood	Moment	Quantile
$\theta$	3	E	3.022	3.026	3.018
		S	0.0626	0.0633	0.0659
$\sigma^2$	0.25	E	0.242	0.229	0.240
		S	0.0438	0.0475	0.0495
E(X)	22.7	E	23.17	23.15	23.09
		S	15.50	15.54	14.78
V(X)	147	E	146.2	140.2	146.3
		S	38.16	40.42	40.07

### 2.3 Confidence-Interval Estimates

For  $\theta$  and  $\sigma^2$

When maximum likelihood estimates of  $\theta$  and  $\sigma^2$  are used, the quantity  $\frac{\hat{\theta} - \theta}{\hat{\sigma}/\sqrt{n}}$  is distributed as t with (n-1) degrees of freedom

and  $\frac{(n-1)\hat{\sigma}^2}{\sigma^2}$  is distributed as  $\chi^2$  with (n - 1) degrees of freedom. Confidence-interval estimates are as follows:

For  $\theta$ , a p percent confidence interval is

$$(\hat{\theta} - t_{p,n-1} \hat{\sigma}/\sqrt{n}, \hat{\theta} + t_{p,n-1} \hat{\sigma}/\sqrt{n})$$

where  $t_{p,n-1}$  is the  $p^{\text{th}}$  percentile of the t distribution with (n-1) degrees of freedom.

For  $\sigma^2$ , a p percent confidence interval is

$$\left( \frac{(n-1)\hat{\sigma}^2}{\chi^2_{p_1, n-1}}, \frac{(n-1)\hat{\sigma}^2}{\chi^2_{p_2, n-1}} \right)$$

where  $\chi^2_{p_1, n-1}$  is the  $\hat{p}_1$ th percentile of the  $\chi^2$  distribution with  $(n-1)$  degrees of freedom and  $p_1 - p_2 = p$

For the moment and quantile estimates, only approximate large-sample confidence intervals may be obtained using the central limit theorem. The general expression for the  $(1 - \alpha)$  percent confidence interval is

$$E \pm Z_{1-\alpha/2} \sqrt{V(E)}$$

where  $Z_{1-\alpha/2}$  is the normal deviate corresponding to the  $(1-\alpha/2)$ th percentile.

Table A-1 gave the theoretical values for  $V(E)$ . Since large samples are assumed, the estimates  $\hat{\theta}$  and  $\hat{\sigma}^2$  can be used in place of  $\theta$  and  $\sigma^2$  when these parameters appear in the equation for  $V(E)$ .

#### For Mean and Variance

Only asymptotic confidence intervals are obtainable for the mean and variance. The sample must be large so that the estimate is assumed to be normal with mean equal to the characteristic being estimated (true mean or true variance) and variance equal to  $V(E)$ . Then a  $(1 - \alpha)$  percent interval is again of the form

$$E \pm Z_{1-\alpha/2} \sqrt{V(E)}$$

The equations for  $V(E)$  are given in Table A-1 and again  $\hat{\theta}$  and  $\hat{\sigma}^2$  are used instead of  $\theta$  and  $\sigma^2$  when such parameters appear in the  $V(E)$  equation.

APPENDIX B  
MULTIPLE-REGRESSION COMPUTER PRINTOUTS FOR  
VARIANCE PREDICTION

EQUATION A(1): VARIANCE OF  $\ln(\text{ACTIVE-REPAIR TIME})$  -- 21 EQUIPMENTS

DATA INPUT

Relative Power (RP)	Signal Data Handling (SD)	Test Concept (TC)	Variance of $\ln$ ART ( $\sigma^2 \ln$ ART)	Square Root of Number of Observations
0.100000E 01	0.300000E 01	0.200000E 01	0.628849E 00	0.400000E 01
0.100000E 01	0.300000E 01	0.250000E 01	0.452929E 00	0.141421E 01
0.100000E 01	0.300000E 01	0.200000E 01	0.233289E 00	0.489042E 01
0.200000E 01	0.500000E 01	0.250000E 01	0.159264E 01	0.282843E 01
0.200000E 01	0.100000E 01	0.200000E 01	0.417316E 00	0.316228E 01
0.200000E 01	0.100000E 01	0.100000E 01	0.250000E 00	0.173205E 01
0.300000E 01	0.200000E 01	0.350000E 01	0.103429E 01	0.591808E 01
0.200000E 01	0.100000E 01	0.350000E 01	0.964324E 00	0.200000E 01
0.200000E 01	0.200000E 01	0.200000E 01	0.190098E 00	0.244949E 01
0.200000E 01	0.100000E 01	0.300000E 01	0.567009E 00	0.519615E 01
0.200000E 01	0.100000E 01	0.300000E 01	0.207936E 00	0.374156E 01
0.200000E 01	0.200000E 01	0.200000E 01	0.801025E 00	0.316228E 01
0.100000E 01	0.200000E 01	0.400000E 01	0.594441E 00	0.244949E 01
0.200000E 01	0.100000E 01	0.300000E 01	0.979844E 00	0.173205E 01
0.300000E 01	0.200000E 01	0.200000E 01	0.107130E 01	0.316228E 01
0.300000E 01	0.200000E 01	0.200000E 01	0.118810E 01	0.489898E 01
0.200000E 01	0.200000E 01	0.200000E 01	0.125888E 01	0.316228E 01
0.100000E 01	0.200000E 01	0.400000E 01	0.521284E 00	0.264575E 01
0.300000E 01	0.300000E 01	0.200000E 01	0.121000E 01	0.640312E 01
0.200000E 01	0.300000E 01	0.200000E 01	0.644809E 00	0.479583E 01
0.200000E 01	0.100000E 01	0.200000E 01	0.345744E 00	0.282843E 01

(continued)

# Equation A(1) (continued)

VARIABLE NO.	MEAN	STD. DEV.
1	2.07143E 00	7.13975E-01
2	2.12956E 00	9.77026E-01
3	2.45672E 00	7.27199E-01
4	7.56218E-01	4.00346E-01

## CORRELATION MATRIX

ROW	COL.	VALUE	ROW	COL.	VALUE	ROW	COL.	VALUE
2	1	-0.09075						
3	1	-0.12240	3	2	-0.22984			
4	1	0.61176	4	2	0.42903	4	3	0.00403

## REGRESSION COEFFS.

INDEX	VALUE	
0	-0.799819E 00	(Constant term)
1	0.385082E 00	(Coefficient of RP)
2	0.221325E 00	(Coefficient of SD)
3	0.116839E 00	(Coefficient of TC)

## PARTIAL CORR. COEFFS., STD. DEV. & T FOR RI

1	0.7540E 00	0.8137E-01	0.4733E 01
2	0.6628E 00	0.6064E-01	0.3650E 01
3	0.3275E 00	0.8175E-01	0.1429E 01

R SQUARED IS 0.652716 R IS 0.807909  
 STD. ERROR IS 0.25590E 00 STD. ERROR SQD. IS 0.65484E-01  
 F IS 0.106504E 02

## GAUSS MULTIPLIERS

ROW	COL.	VALUE	ROW	COL.	VALUE	ROW	COL.	VALUE
1	1	0.1011E 00	1	2	0.9274E-07	1	3	0.1501E-01
2	2	0.5615E-01	2	3	0.1845E-11	3	3	0.1021E 00

Equation A(1) (continued)

TEST INDEPENDENT VARIABLES FOR SIGNIFICANCE

1 ST VAR. IS 1 R<sup>2</sup> = 0.374249 F = 11.364 CRIT. VALUE IS 1.760  
 2 ND VAR. IS 2 R<sup>2</sup> = 0.610986 F = 10.954 CRIT. VALUE IS 1.770  
 3 RD VAR. IS 3 R<sup>2</sup> = 0.652716 F = 2.043 CRIT. VALUE IS 1.780

ALL VARIABLES SIGNIFICANT. SELECTION ORDER IS 1 2 3

ACTUAL VS. PREDICTED RESULTS

OBSERVATION	ACTUAL	PREDICTED	DEVIATION	WEIGHTED DEV.
1	0.62885E 00	0.48292E 00	0.14593E 00	0.16902E 00
2	0.45293E 00	0.54134E 00	-0.88408E-01	-0.36202E-01
3	0.23329E 00	0.48292E 00	-0.24963E 00	-0.33903E 00
4	0.15026E 01	0.13591E 01	0.22357E 00	0.18310E 00
5	0.41732E 00	0.42535E 00	-0.80328E-02	-0.73552E-02
6	0.25000E 00	0.30851E 00	-0.58509E-01	-0.29344E-01
7	0.10343E 01	0.12070E 01	-0.17272E 00	-0.29508E 00
8	0.96432E 00	0.60061E 00	0.36372E 00	0.21063E 00
9	0.19010E 00	0.64667E 00	-0.45658E 00	-0.32383E 00
10	0.56701E 00	0.54219E 00	0.24821E-01	0.37345E-01
11	0.20794E 00	0.54219E 00	-0.33425E 00	-0.36213E 00
12	0.80103E 00	0.64667E 00	0.15435E 00	0.14133E 00
13	0.59444E 00	0.49527E 00	0.99170E-01	0.70337E-01
14	0.87984E 00	0.54219E 00	0.33766E 00	0.16934E 00
15	0.10733E 01	0.10318E 01	0.41544E-01	0.38040E-01
16	0.11881E 01	0.10318E 01	0.15634E 00	0.22178E 00
17	0.12589E 01	0.64667E 00	0.61221E 00	0.58792E 00
18	0.52128E 00	0.49527E 00	0.26013E-01	0.19928E-01
19	0.12100E 01	0.12531E 01	-0.43081E-01	-0.79874E-01
20	0.64401E 00	0.86800E 00	-0.22319E 00	-0.30993E 00
21	0.34574E 00	0.42535E 00	-0.79605E-01	-0.65195E-01
STD. DEV. OF DEV. IS 0.13770E 00, VAR. IS 0.18961E-01				
A/G. DEV. IS -0.37982E-07				

EQUATION A(2): VARIANCE OF  $\ln(\text{ACTIVE-REPAIR TIME})$  -- 13 EQUIPMENTS

DATA INPUT

Maintenance Complexity (MC)	Relative Power (RP)	Efficiency of Information Transmission (ET)	Variance of $\ln \text{ART}$ ( $\sigma^2 \ln \text{ART}$ )	Square Root of Number of Observations
0.120000E 02	0.100000E 01	0.201000E 00	0.452929E 00	0.141421E 01
0.160000E 02	0.100000E 01	0.479000E 00	0.233289E 00	0.489042E 01
0.180000E 02	0.200000E 01	0.156000E 00	0.417316E 00	0.316228E 01
0.140000E 02	0.200000E 01	0.790000E-01	0.964324E 00	0.200000E 01
0.120000E 02	0.200000E 01	0.472000E 00	0.190096E 00	0.244949E 01
0.140000E 02	0.200000E 01	0.460000E 00	0.567009E 00	0.519615E 01
0.120000E 02	0.200000E 01	0.477000E 00	0.207936E 00	0.374166E 01
0.140000E 02	0.200000E 01	0.485000E 00	0.801025E 00	0.316228E 01
0.250000E 02	0.100000E 01	0.438000E 00	0.594441E 00	0.244949E 01
0.170000E 02	0.200000E 01	0.840000E-01	0.879844E 00	0.173205E 01
0.310000E 02	0.300000E 01	0.327000E 00	0.121000E 01	0.640312E 01
0.230000E 02	0.200000E 01	0.482000E 00	0.644809E 00	0.479583E 01
0.240000E 02	0.200000E 01	0.722000E 00	0.345744E 00	0.282843E 01

(continued)

Equation A(2) (continued)

VARIABLE NO.

MEAN

STD. DEV.

1	1.89903E 01	6.73171E 09
2	1.95114E 00	6.04539E 01
3	4.02337E 01	1.63919E 01
4	4.02020E 01	3.65924E 01

CORRELATION MATRIX

ROW	COL.	VALUE	ROW	COL.	VALUE	ROW	COL.	VALUE
2	1	0.52042						
3	1	0.01715	3	2	-0.16006			
4	1	0.65104	4	2	0.68856	4	3	-0.42610

REGRESSION COEFFS.

INDEX	VALUE	
0	0.365817E-01	(Constant Term)
1	0.233132E-01	(Coefficient of MC)
2	0.224846E 00	(Coefficient of RP)
3	-0.784817E 00	(Coefficient of ET)

PARTIAL CORR. COEFFS., STD. DEV. & T FOR R1

1	0.5909E 00	0.1061E-01	0.2197E 01
2	0.5308E 00	0.1197E 00	0.1879E 01
3	-0.5702E 00	0.3769E 00	-0.2082E 01

R SQUARED IS 0.724381    R IS 0.851106  
 STD. ERROR IS 0.20971E 00    STD. ERROR SQD. IS 0.43977E-01  
 F IS 0.788458E 01

GAUSS MULTIPLIERS

ROW	COL.	VALUE	ROW	COL.	VALUE	ROW	COL.	VALUE
1	1	0.2360E-02	1	2	-0.1531E-01	1	3	-0.1105E-01
2	2	0.3256E 00	2	3	0.2043E 00	3	3	0.3231E 01

Equation A(2) (continued)

TEST INDEPENDENT VARIABLES FOR SIGNIFICANCE

1 ST VAR.IS 2 R<sup>2</sup> = 0.474119 F = 9.917 CRIT. VALUE IS 1.860  
 2 ND VAR.IS 1 R<sup>2</sup> = 0.591614 F = 2.877 CRIT. VALUE IS 1.880  
 3 RD VAR.IS 3 R<sup>2</sup> = 0.724381 F = 4.335 CRIT. VALUE IS 1.910

ALL VARIABLES SIGNIFICANT. SELECTION ORDER IS 2 1 3

ACTUAL VS. PREDICTED RESULTS

OBSERVATION	ACTUAL	PREDICTED	DEVIATION	WEIGHTED DEV.
1	0.45293E 00	0.38344E 00	0.69490E-01	0.29019E-01
2	0.23329E 00	0.25851E 00	-0.25223E-01	-0.34935E-01
3	0.41732E 00	0.78348E 00	-0.36616E 00	-0.34191E 00
4	0.96432E 00	0.75066E 00	0.21367E 00	0.12618E 00
5	0.19010E 00	0.39560E 00	-0.20550E 00	-0.14864E 00
6	0.56701E 00	0.45164E 00	0.11537E 00	0.17701E 00
7	0.20794E 00	0.39168E 00	-0.18374E 00	-0.20300E 00
8	0.80103E 00	0.43202E 00	0.36900E 00	0.34456E 00
9	0.59444E 00	0.50051E 00	0.93932E-01	0.67941E-01
10	0.87984E 00	0.81667E 00	0.63170E-01	0.32308E 01
11	0.12100E 01	0.11772E 01	0.32805E-01	0.62021E 01
12	0.64481E 00	0.64470E 00	0.61223E-03	0.86700E-03
13	0.34574E 00	0.47915E 00	-0.13341E 00	-0.11142E 00
STD. DEV. OF DEV. IS 0.11395E 00, VAR. IS 0.12986E-01				
AVG. DEV. IS -0.11514E-07				

EQUATION B: VARIANCE OF  $\Delta n$  (ACTIVE MAN-HOURS) -- 21 EQUIPMENTS

DATA INPUT

Relative Power (RP)	Signal Data Handling (SD)	Test Concept (TC)	Variance of $\Delta n$ MH ( $\sigma^2 \Delta n$ MH)	Square Root of Number of Observations
0.100000E 01	C.300000E 01	0.200000E 01	0.680625E 00	0.400000E 01
0.100000E 01	0.300000E 01	0.250000E 01	0.135257E 01	0.141421E 01
0.100000E 01	0.300000E 01	0.200000E 01	0.364816E 00	0.469042E 01
0.200000E 01	0.500000E 01	0.250000E 01	0.149573E 01	0.282863E 01
0.200000E 01	0.100000E 01	0.200000E 01	0.546121E 00	0.316228E 01
0.200000E 01	0.100000E 01	0.100000E 01	0.425104E 00	0.173205E 01
0.300000E 01	0.200000E 01	0.350000E 01	0.159770E 01	0.591608E 01
0.200000E 01	0.100000E 01	0.350000E 01	0.155003E 01	0.200000E 01
0.200000E 01	0.200000E 01	0.200000E 01	0.378225E 00	0.244949E 01
0.200000E 01	0.100000E 01	0.300000E 01	0.732736E 00	0.519615E 01
0.200000E 01	0.100000E 01	0.300000E 01	0.802616E 00	0.374166E 01
0.200000E 01	0.200000E 01	0.200000E 01	0.105268E 01	0.316228E 01
0.100000E 01	0.200000E 01	0.400000E 01	0.594441E 00	0.244949E 01
0.200000E 01	0.100000E 01	0.300000E 01	0.753424E 00	0.173205E 01
0.300000E 01	0.200000E 01	0.200000E 01	C.154008E 01	0.316228E 01
0.300000E 01	0.200000E 01	0.200000E 01	0.163840E 01	0.489898E 01
0.200000E 01	0.200000E 01	0.200000E 01	0.155501E 01	0.331662E 01
0.100000E 01	0.200000E 01	0.400000E 01	0.412164E 00	0.264575E 01
0.300000E 01	0.300000E 01	0.200000E 01	0.127013E 01	0.640312E 01
0.200000E 01	0.300000E 01	0.200000E 01	0.109412E 01	0.479383E 01
0.200000E 01	0.100000E 01	0.200000E 01	0.600625E 00	0.282863E 01

(continued)

(Equation B (continued))

VARIABLE NO.	MEAN	STD. DEV.
1	2.07143E 00	7.13976E-01
2	2.12956E 00	9.77026E-01
3	2.45672E 00	7.27199E-01
4	1.01541E 00	4.64341E-01

# CORRELATION MATRIX

ROW	COL.	VALUE	ROW	COL.	VALUE	ROW	COL.	VALUE
2	1	-0.09075						
3	1	-0.12240	3	2	-0.22984			
4	1	0.71304	4	2	0.23930	4	3	0.03167

# REGRESSION COEFFS.

INDEX	VALUE	
0	-0.711014E 00	(Constant term)
1	0.501373E 00	(Coefficient of PP)
2	0.169711E 00	(Coefficient of SD)
3	0.132883E 00	(Coefficient of TC)

# PARTIAL CORR. COEFFS., STD. DEV. & T FOR R:

1	0.7853E 00	0.9566E-01	0.5230E 01
2	0.4993E 00	0.7143E-01	0.2376E 01
3	0.3174E 00	0.9630E-01	0.1380E 01

R SQUARED IS 0.641736 R IS 0.801084  
 STD. ERROR IS 0.30146E 00 STD. ERROR SQD. IS 0.90878E-01  
 F IS 0.101503E 02

# GAUSS MULTIPLIERS

ROW	COL.	VALUE	ROW	COL.	VALUE	ROW	COL.	VALUE
1	1	0.1011E 00	1	2	0.9274E-02	1	3	0.1501E-01
2	2	0.5615E-01	2	3	0.1845E-01	3	3	0.1021E 00

(continued)

Equation B (continued)

TEST INDEPENDENT VARIABLES FOR SIGNIFICANCE

1 ST VAR. IS 1 R<sup>2</sup> = 0.508423 F = 19.651 CRIT. VALUE IS 1.760  
 2 ND VAR. IS 2 R<sup>2</sup> = 0.601611 F = 5.210 CRIT. VALUE IS 1.770  
 3 RD VAR. IS 3 R<sup>2</sup> = 0.641736 F = 1.904 CRIT. VALUE IS 1.780

ALL VARIABLES SIGNIFICANT. SELECTION ORDER IS 1 2 3

ACTUAL VS. PREDICTED RESULTS

OBSERVATION	ACTUAL	PREDICTED	DEVIATION	WEIGHTED DEV.
1	0.68063E 00	0.56526E 00	0.11537E 00	0.13362E 00
2	0.13526E 01	0.63170E 00	0.72087E 00	0.29519E 00
3	0.36482E 00	0.56526E 00	-0.20044E 00	-0.27223E 00
4	0.14957E 01	0.14725E 01	0.23233E-01	0.19027E-01
5	0.54612E 00	0.72721E 00	-0.18109E 00	-0.16561E 00
6	0.42310E 00	0.59433E 00	-0.16922E 00	-0.84869E-01
7	0.15977E 01	0.15976E 01	0.80730E-04	0.13829E-03
8	0.15500E 01	0.92654E 00	0.62349E 00	0.36107E 00
9	0.37823E 00	0.89692E 00	-0.51870E 00	-0.36789E 00
10	0.73274E 00	0.86009E 00	-0.12736E 00	-0.19162E 00
11	0.80282E 00	0.86009E 00	-0.05727E-01	-0.62055E-01
12	0.10527E 01	0.89692E 00	0.15576E 00	0.14262E 00
13	0.59444E 00	0.66132E 00	-0.06687E-01	-0.47431E-01
14	0.75342E 00	0.86009E 00	-0.10667E 00	-0.53497E-01
15	0.15401E 01	0.13983E 01	0.14179E 00	0.12983E 00
16	0.16384E 01	0.13983E 01	0.24011E 00	0.34059E 00
17	0.15550E 01	0.89692E 00	0.65809E 00	0.63199E 00
18	0.41216E 00	0.66132E 00	-0.24915E 00	-0.19087E 00
19	0.12701E 01	0.15680E 01	-0.29788E 00	-0.59227E 00
20	0.10941E 01	0.10666E 01	0.27487E-01	0.38170E-01
21	0.60063E 00	0.72721E 00	-0.12659E 00	-0.10367E 00
STD. DEV. OF DEV. IS 0.16222E 00, VAR. IS 0.26314E-01				
AVG. DEV. IS 0.28607E-06				

UNCLASSIFIED

Security Classification	
DOCUMENT CONTROL DATA - R & D	
(Security classification of title, body or abstract and indexing annotation must be entered when the overall report is classified)	
1. ORIGINATING ACTIVITY (Corporate author) ARINC Research Corporation 2551 Riva Road Annapolis MD 21401	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP
3. REPORT TITLE MAINTAINABILITY PREDICTION AND DEMONSTRATION TECHNIQUES Vol I Development of Prediction Techniques Vol II Maintainability Demonstration	
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Interim Reports Dec 67 - Dec 69	
5. AUTHOR(S) (First name, middle initial, last name) Vol I Grafton H. Griswold Seymour Bishop Bernard L. Rettérer John Martin Angelo C. Mitaopoulos Vol II Harold S. Balaban	
6. REPORT DATE January 1970	7a. TOTAL NO. OF PAGES Vol I 334 Vol II 255 7b. NO. OF REFS Vol I 10 Vol II 15
8a. CONTRACT OR GRANT NO. F30602-68-C-0047	8b. ORIGINATOR'S REPORT NUMBER(S) 573-01-1-992
9. PROJECT NO. 5519 a. Task No. 551901 d.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned to this report) RADC-TR-69-356, Vols I and II
10. DISTRIBUTION STATEMENT This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of RADC (EMNRR), GARD, NY 13440.	
11. SUPPLEMENTARY NOTES PROJECT ENGINEER Eugene Fiorentino	12. SPONSORING MILITARY ACTIVITY Rome Air Development Center (EMNRR) Griffiss Air Force Base, New York 13440
13. ABSTRACT Maintainability prediction is a design tool that can be used to meet desired maintainability goals during system development and design. Volume I contains the results of a study to develop improved maintainability prediction techniques for use on all major classes of Air Force electronic systems at the equipment level of maintenance. The techniques were developed for application from concept formulation through the detail design phases of system development. A data collection program was conducted to acquire detailed information on maintenance actions performed in the field. The data were analyzed to determine relationships between maintenance design variables and maintenance time. Prediction models were developed for corrective and preventive maintenance. Predictive relationships between maintenance time and maintenance manhours were also developed. A data base was established for use in future investigation of maintainability.  Maintainability demonstration is a testing procedure for assuring the acquisition of equipment and systems which meet specified numerical maintainability requirements. Volume II contains the results of a study to develop improved maintainability demonstration procedures for all major classes of Air Force equipment. An industry and government-wide survey was conducted to provide insight into the current status of maintainability demonstration and to initiate research into the technical and administrative aspects of demonstration. Specific recommendations and guidelines were developed for maintainability index selection, maintenance task sampling procedures, statistical maintainability demonstration plans, and for test planning and administration. The use of prior information for specifying numerical requirements, designing (Over)	

DD FORM 1 NOV 66 1473

UNCLASSIFIED  
Security Classification

UNCLASSIFIED  
Security Classification

1a.	KEY WORDS	LINK A		LINK B		LINK C	
		ROLE	WT	ROLE	WT	ROLE	WT
	Maintainability Electronic Equipment Statistical Techniques						
	<u>Abstract Continued</u> statistical sampling procedures, developing test criteria, and applying Bayesian tests was also investigated and the results are presented in the report.						

UNCLASSIFIED  
Security Classification

APUC-Griffins AFB TX 18 May 70-57